APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

April 2020
Vol. 35 No. 4
ISSN 1054-4887

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

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

APRIL 2020 REVIEWERS: REGULAR PAPERS

Maan Abdulwahid
Ramin Aghajafari
Ugur Alkasi
Amrita Bal
Mehmet Belen
Guan-Yu Chen
Tijana Dimitrijevic
Yingsong Li
Wen-Jiao Liao
Yang Liu
Ehsan Mostafapour
Satyanarayana Murthy
Mohammd Pourbagher
Jagdishkumar Rathod
Daniele Romano
Hai Rong
Kamalesh Sainath
Kenedy Santos
Masoud Sarabi
S. Saravanan
T. Shanmuganantham
Singaravelan Shanmugasundaram
Ramy Sharafeldin
Ashish Singh
Yogesh Thakare
Koushick V.
Ioannis Vardiambasis
Chakravarthy Vedula
Qi Wu
Salah Yahya
Wenhua Yu
Lei Zhao

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

Vol. 35 No. 4 April 2020

TABLE OF CONTENTS – REGULAR PAPERS

An FDTD TF-SF Boundary on Face-Centered Cubic Grids

Lijuan Shi [1], Lixia Yang [2], and Wei Feng [1]

[1] Department of Physics
Jiangsu University, Zhenjiang, 212013, China
juan-online@163.com, wfeng@ujs.edu.cn

[2] Department of Communication Engineering
Jiangsu University, Zhenjiang, 212013, China
lixiayang@yeah.net

Abstract ─ The implementation of total-field/scattered-field (TF-SF) boundary for finite-difference time-domain (FDTD) method was proposed based on face-centered cubic (FCC) grids. On the basis of the arrangement of fields components in the FCC voxel, the update equations for the electric and magnetic field components of the TF-SF boundary are derived. The resonant frequency of a cavity is calculated to illustrate that the FCC grid scheme is more accurate as compared to the equivalent Yee grid method. Then, by simulating the amplitude distribution of the vertical and oblique incident plane waves in the whole calculation region, the effectiveness of the TF-SF boundary is verified numerically.

Index Terms ─ Face-centered cubic grids, finite-difference time-domain method, total-field/scattered-field boundary.

I. INTRODUCTION

The finite-difference time-domain (FDTD) method is widely applied in many electromagnetic problems for its instinctive structure and simple implementation [1-6]. It is well known that the algorithmic structure of the standard FDTD method is based on Cartesian (Yee) grid [7,8]. However, due to the selection of computational molecule and grid, this Yee-based method has one drawback of anisotropy as described by the numerical dispersive equation. Thus, the wave-front distortions which would not appear in the analytic domain can be demonstrated in numerical simulations. One effective way of improving the isotropy of FDTD scheme is to use alternatives to the standard Cartesian grid [9-11]. Potter proposed the face-centered cubic (FCC) lattices for the discretization of Maxwell's equation [12]. The comparison of dispersion and stability analysis of this method with an equivalent Cartesian method has shown that the isotropy has been significantly improved, and the stability criterion has also been much relaxed than the Yee cell ($\Delta t|_{FCC} \approx 1.37473\Delta t|_{Yee}$), resulting in further computational savings. As a result, the FCC grid can be considered as an attractive alternative tesselation of 3D space for Maxwell's equations.

In order to apply the FDTD method based upon FCC grids to electromagnetic scattering problems of complex mediums, efficient total-field/scattered-field (TF-SF) boundary need to be introduced into the computation region [13,14]. The TF-SF formulation divides the entire FDTD lattice into two distinct regions. Incident and scattered fields exist in the total-field region, and scattered fields are only limited in the scattered-field region [15,16]. Failure to do so can lead to high levels of field leakage errors across the TF-SF boundary. In this paper, we present a detailed explanation of the implementation of TF-SF boundary for FDTD method based upon FCC grids.

The organization of this paper is as follows: Section 2 simply introduces the FCC grid structure. In Section 3, the update equations for the electric and magnetic field components of the TF-SF boundary are derived. The effectiveness of the proposed scheme is numerically verified in Section 4. Section 5 are the conclusions.

II. DESCRIPTION OF THE FCC GRID STRUCTURE

The face-centered cubic (FCC) grid is a spatial extension of hexagonal grid. The arrangement of electric and magnetic fields in the FCC voxel is shown in Fig. 1, which consists of lattice points lying at each corner and in the center of all six faces [5]. Two FCC grids should be considered for discretizing Maxwell's equations: one indicates the electric field components, and the other hosts the magnetic components. Different from the conventional Cartesian (Yee) grid, both electric field components and magnetic field components are collocated (each lattice point denote all three components) in FCC grids, whereas magnetic and electric fields are

Submitted On: August 6, 2019
Accepted On: April 12, 2020

staggered. It is obviously seen that there are four unique electric field vectors and four unique magnetic field vectors in the FCC voxel: located at the closest corners, as well as lying in the center of the xy-face, yz-face and xz-face. For the purposes of indexing and analysis, the side lengths of the FCC voxel are assumed as $\Delta x, \Delta y$ and Δz, the indies with a given voxel are set as $(i, j, k, n) = (i\Delta x, j\Delta y, k\Delta z, n\Delta t)$, where Δt is the time step. The magnetic field grid is displaced from the electric field grid along the direction of the voxel main diagonal by the displacement vector $d = (\Delta x/4, \Delta y/4, \Delta z/4)$.

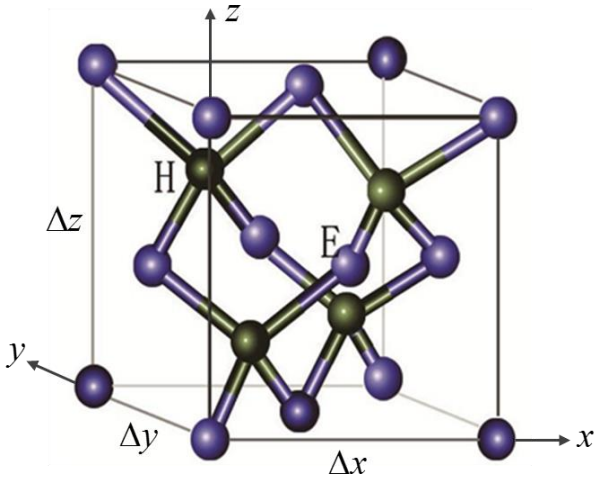


Fig. 1. Electric and magnetic field lattice points in the FCC voxel.

According to the arrangement of fields components in the FCC voxel, the tetrahedral structure of electric fields surrounding a given magnetic field is shown in Fig. 2 (a). The field update equations at the tetrahedron center are obtained using these nearest neighbors. Projecting this tetrahedral structure onto the rectangular coordinate axes, the displacement vectors that represent the location of these electric field components with reference to the magnetic field location, are $\boldsymbol{e}_1 = \left(-\frac{\Delta x}{4}, \frac{\Delta y}{4}, \frac{\Delta z}{4}\right)$, $\boldsymbol{e}_2 = \left(\frac{\Delta x}{4}, -\frac{\Delta y}{4}, \frac{\Delta z}{4}\right)$, $\boldsymbol{e}_3 = \left(\frac{\Delta x}{4}, \frac{\Delta y}{4}, -\right)$, and $\boldsymbol{e}_4 = \left(-\frac{\Delta x}{4}, -\frac{\Delta y}{4}, -\frac{\Delta z}{4}\right)$, respectively. On the contrary, in Fig. 2 (b) the appropriate displacement vectors presenting the magnetic field locations neighboring the electric field, are $\boldsymbol{h}_1 = \left(\frac{\Delta x}{4}, \frac{\Delta y}{4}, \frac{\Delta z}{4}\right)$, $\boldsymbol{h}_2 = \left(-\frac{\Delta x}{4}, -\frac{\Delta y}{4}, \frac{\Delta z}{4}\right)$, $\boldsymbol{h}_3 = \left(-\frac{\Delta x}{4}, \frac{\Delta y}{4}, -\frac{\Delta z}{4}\right)$ and $\boldsymbol{h}_4 = \left(\frac{\Delta x}{4}, -\frac{\Delta y}{4}, -\frac{\Delta z}{4}\right)$, respectively.

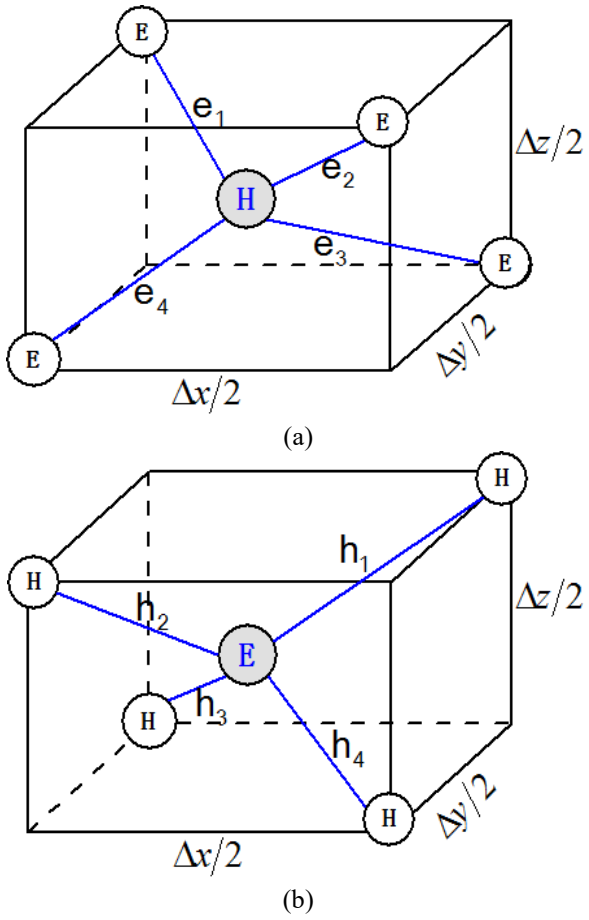


Fig. 2. (a) The organization of the electric fields neighboring a specified magnetic field point; and (b) the organization of the magnetic fields neighboring a specified electric field point.

III. TF-SF BOUNDARY IMPLEMENTATION ON FCC GRIDS

As illustrated in Fig. 3 (a), the field update points around the TF-SF boundary for the FDTD method on FCC grids are different from those on the Yee scheme. The field update points on the TF-SF boundary are assumed as electric field grids, and the grid points displace $\Delta_m/4$ ($m \in x, y, z$) away from the TF/SF boundary are set as magnetic field grids. Here, we take the upper boundary as an example. The derivations for other five boundary proceeds in a similar way. In Fig. 3 (b), the filled circle ② at the upper boundary indicates the electric field grid located at the corners of FCC grids, and circles ① and ③ indicate electric field grid lying in the center of the face of an FCC grid. The hollow circles ④-⑦ indicate the neighboring magnetic field grids on the exterior boundary located $\Delta_z/4$ distance away from the upper boundary.

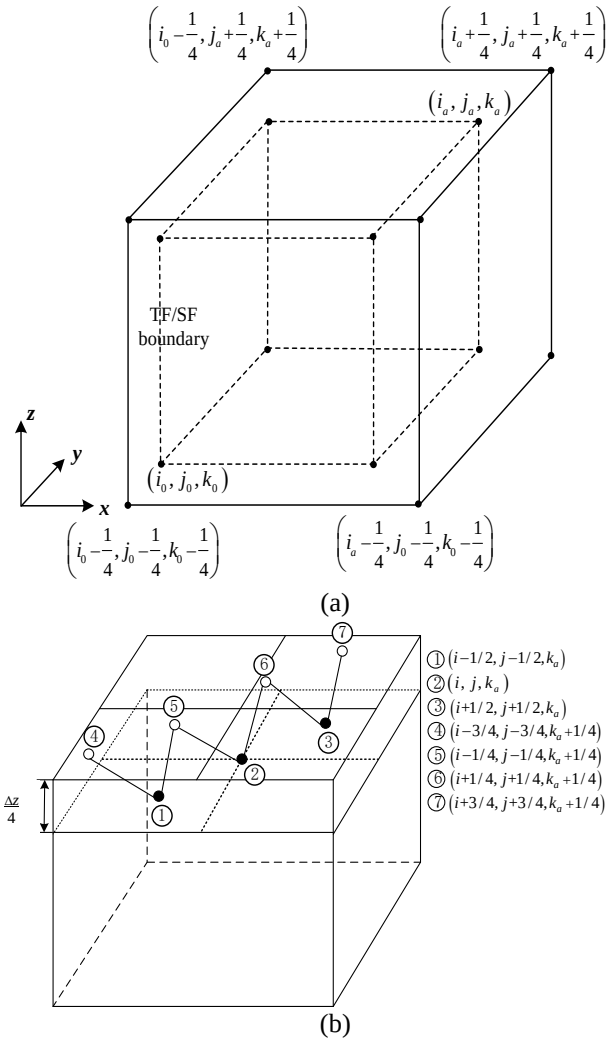


Fig. 3. (a)TF-SF field boundary of FCC-FDTD method; and (b) arrangement of fields components at the upper boundary.

A. Update equations of electric field components

It is obvious that the computations of the electric field components on the TF-SF boundary are related to the magnetic field components on the exterior boundary which belong to the SF field. Therefore, when the incident wave is introduced into TF region by the TF-SF boundary, the iteration equations of electric field components are required to add the incident wave field value at the corresponding time step and the grid [1]. Thus, the update equations of electric field components at the upper boundary ($z = z_a$) can be obtained as below:

(1) Electric field grid located at the corners of FCC grids:

$$E_x\big|_{i,j,k_a}^{n+1} = \left\{E_x\big|_{i,j,k_a}^{n+1}\right\}_{FCC} + \frac{\Delta t}{\Delta y\cdot\varepsilon}\left[H_{z,i}\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} - H_{z,i}\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right] - \frac{\Delta t}{\Delta z\cdot\varepsilon}\left[H_{y,i}\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} + H_{y,i}\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right], \quad (1)$$

where $\left\{E_x\big|_{i,j,k_a}^{n+1}\right\}_{FCC}$ is update equation for electric component E_x at location (i, j, k_a) based upon FCC grids:

$$\left\{E_x\big|_{i,j,k_a}^{n+1}\right\}_{FCC} = E_x\big|_{i,j,k_a}^{n} + \frac{\Delta t}{\Delta y\cdot\varepsilon}\left[H_z\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} + H_z\big|_{i-\frac{1}{4},j+\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2} - H_z\big|_{i+\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} - H_z\big|_{i+\frac{1}{4},j-\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2}\right] - \frac{\Delta t}{\Delta z\cdot\varepsilon}\left[H_y\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} + H_y\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} - H_y\big|_{i-\frac{1}{4},j+\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2} - H_y\big|_{i+\frac{1}{4},j-\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2}\right]. \quad (2)$$

Similarly, the update equations can be obtained for $E_y\big|_{i,j,k_a}^{n+1}$ and $E_z\big|_{i,j,k_a}^{n+1}$,

$$E_y\big|_{i,j,k_a}^{n+1} = \left\{E_y\big|_{i,j,k_a}^{n+1}\right\}_{FCC} + \frac{\Delta t}{\Delta z\cdot\varepsilon}\left[H_{x,i}\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} + H_{x,i}\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right] - \frac{\Delta t}{\Delta x\cdot\varepsilon}\left[H_{z,i}\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} - H_{z,i}\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right], \quad (3)$$

$$E_z\big|_{i,j,k_a}^{n+1} = \left\{E_z\big|_{i,j,k_a}^{n+1}\right\}_{FCC} + \frac{\Delta t}{\Delta x\cdot\varepsilon}\left[H_{y,i}\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} - H_{y,i}\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right] - \frac{\Delta t}{\Delta y\cdot\varepsilon}\left[H_{x,i}\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} + H_{x,i}\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right], \quad (4)$$

where

$$\left\{E_y\big|_{i,j,k_a}^{n+1}\right\}_{FCC} = E_y\big|_{i,j,k_a}^{n} + \frac{\Delta t}{\Delta z\cdot\varepsilon}\left[H_x\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} + H_x\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} - H_x\big|_{i-\frac{1}{4},j+\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2} - H_x\big|_{i+\frac{1}{4},j-\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2}\right] - \frac{\Delta t}{\Delta x\cdot\varepsilon}\left[H_z\big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} + H_z\big|_{i+\frac{1}{4},j-\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2} - H_z\big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2} - H_z\big|_{i-\frac{1}{4},j+\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2}\right], \quad (5)$$

$$\left\{E_z\Big|_{i,j,k_a}^{n+1}\right\}_{FCC}=E_z\Big|_{i,j,k_a}^{n}+\frac{\Delta t}{\Delta x\cdot\varepsilon}\left[H_y\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}+H_y\Big|_{i+\frac{1}{4},j-\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2}-H_y\Big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}-H_z\Big|_{i-\frac{1}{4},j+\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2}\right]-\frac{\Delta t}{\Delta y\cdot\varepsilon}\left[H_x\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}+H_x\Big|_{i-\frac{1}{4},j+\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2}-H_x\Big|_{i-\frac{1}{4},j-\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}-H_x\Big|_{i+\frac{1}{4},j-\frac{1}{4},k_a-\frac{1}{4}}^{n+1/2}\right]. \quad (6)$$

In addition, for the electric field components located at four edges of the upper boundary, they also belong to the corners of FCC grids, and can be derived in the same way.

(2) Electric field grid lying in the center of the face of FCC grids:

$$E_x\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}=\left\{E_x\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}\right\}_{FCC}+\frac{\Delta t}{\Delta y\cdot\varepsilon}\left[H_{z,i}\Big|_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}^{n+1/2}-H_{z,i}\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right]-\frac{\Delta t}{\Delta z\cdot\varepsilon}\left[H_{y,i}\Big|_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}^{n+1/2}+H_{y,i}\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right], \quad (7)$$

$$E_y\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}=\left\{E_y\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}\right\}_{FCC}+\frac{\Delta t}{\Delta z\cdot\varepsilon}\left[H_{x,i}\Big|_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}^{n+1/2}-H_{x,i}\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right]-\frac{\Delta t}{\Delta x\cdot\varepsilon}\left[H_{z,i}\Big|_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}^{n+1/2}+H_{z,i}\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right], \quad (8)$$

$$E_z\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}=\left\{E_z\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}\right\}_{FCC}+\frac{\Delta t}{\Delta x\cdot\varepsilon}\left[H_{y,i}\Big|_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}^{n+1/2}-H_{y,i}\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right]-\frac{\Delta t}{\Delta y\cdot\varepsilon}\left[H_{x,i}\Big|_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}^{n+1/2}+H_{x,i}\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right]. \quad (9)$$

Here, $\left\{E_y\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}\right\}_{FCC}$, $\left\{E_y\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}\right\}_{FCC}$ and $\left\{E_y\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n+1}\right\}_{FCC}$ in equations (7-9) are the update equations of electric field components at node $\left\{i+\frac{1}{2},j_0,k+\frac{1}{2}\right\}$, which can be easily obtained by replacing $\{i,j,k_a\}$ in equations (2), (5) and (6) with $\left\{i+\frac{1}{2},j_0,k+\frac{1}{2}\right\}$, respectively .

B. Update equations of magnetic field components

In Fig. 3, it is also seen that the computations of magnetic field components on the exterior boundary located $\Delta_z/4$ distance away from the upper boundary are related to the electric field components on the TF-SF boundary, and therefore the iteration equations of magnetic field components are required to subtract the incident wave field value at the corresponding the grid on the TF-SF boundary. Then, the update equations of magnetic field components can be represented as below:

(1) Magnetic field grid located at the corners of FCC grids:

$$H_x\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}=\left\{H_x\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right\}_{FCC}+\frac{\Delta t}{\Delta z\cdot\mu}\left[E_{y,i}\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n}+E_{y,i}\Big|_{i,j,k_a}^{n}\right]+\frac{\Delta t}{\Delta y\cdot\mu}\left[E_{z,i}\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n}-E_{z,i}\Big|_{i,j,k_a}^{n}\right], \quad (10)$$

$$H_y\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}=\left\{H_y\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right\}_{FCC}-\frac{\Delta t}{\Delta x\cdot\mu}\left[E_{z,i}\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n}-E_{z,i}\Big|_{i,j,k_a}^{n}\right]-\frac{\Delta t}{\Delta z\cdot\mu}\left[E_{x,i}\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n}+E_{x,i}\Big|_{i,j,k_a}^{n}\right], \quad (11)$$

$$H_z\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}=\left\{H_z\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right\}_{FCC}-\frac{\Delta t}{\Delta y\cdot\mu}\left[E_{x,i}\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n}-E_{x,i}\Big|_{i,j,k_a}^{n}\right]+\frac{\Delta t}{\Delta x\cdot\mu}\left[E_{y,i}\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n}-E_{y,i}\Big|_{i,j,k_a}^{n}\right], \quad (12)$$

where $\left\{H_x\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right\}_{FCC}$, $\left\{H_y\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right\}_{FCC}$ and $\left\{H_z\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right\}_{FCC}$ are the update equations of magnetic field components at location $\left\{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}\right\}$ based upon FCC grids:

$$\left\{H_x\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n+1/2}\right\}_{FCC}=H_x\Big|_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}^{n-1/2}+\frac{\Delta t}{\Delta z\cdot\mu}\left[E_y\Big|_{i,j+\frac{1}{2},k_a+\frac{1}{2}}^{n}+E_y\Big|_{i+\frac{1}{2},j,k_a+\frac{1}{2}}^{n}-E_y\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n}-E_y\Big|_{i,j,k_a}^{n}\right]-\frac{\Delta t}{\Delta y\cdot\mu}\left[E_z\Big|_{i,j+\frac{1}{2},k_a+\frac{1}{2}}^{n}+E_z\Big|_{i+\frac{1}{2},j+\frac{1}{2},k_a}^{n}-E_z\Big|_{i+\frac{1}{2},j,k_a+\frac{1}{2}}^{n}-E_z\Big|_{i,j,k_a}^{n}\right], \quad (13)$$

$$\left\{H_y\Big|^{n+1/2}_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}\right\}_{FCC}=H_y\Big|^{n-1/2}_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}+\frac{\Delta t}{\Delta x\cdot\mu}\left[E_z\Big|^n_{i+\frac{1}{2},j,k_a+\frac{1}{2}}+E_z\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}-E_z\Big|^n_{i,j+\frac{1}{2},k_a+\frac{1}{2}}-E_z\Big|^n_{i,j,k_a}\right]-\frac{\Delta t}{\Delta z\cdot\mu}\left[E_x\Big|^n_{i,j+\frac{1}{2},k_a+\frac{1}{2}}+E_x\Big|^n_{i+\frac{1}{2},j,k_a+\frac{1}{2}}-E_x\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}-E_x\Big|^n_{i,j,k_a}\right], \quad (14)$$

$$\left\{H_z\Big|^{n+1/2}_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}\right\}_{FCC}=H_z\Big|^{n-1/2}_{i+\frac{1}{4},j+\frac{1}{4},k_a+\frac{1}{4}}+\frac{\Delta t}{\Delta y\cdot\mu}\left[E_x\Big|^n_{i,j+\frac{1}{2},k_a+\frac{1}{2}}+E_x\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}-E_x\Big|^n_{i+\frac{1}{2},j,k_a+\frac{1}{2}}-E_x\Big|^n_{i,j,k_a}\right]-\frac{\Delta t}{\Delta x\cdot\mu}\left[E_y\Big|^n_{i+\frac{1}{2},j,k_a+\frac{1}{2}}+E_y\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}-E_y\Big|^n_{i,j+\frac{1}{2},k_a+\frac{1}{2}}-E_y\Big|^n_{i,j,k_a}\right]. \quad (15)$$

(2) Magnetic field grid lying in the center of the face of FCC grids:

$$H_x\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}=\left\{H_x\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}\right\}_{FCC}+\frac{\Delta t}{\Delta z\cdot\mu}\left[E_{y,i}\Big|^n_{i+1,j+1,k_a}+E_{y,i}\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}\right]+\frac{\Delta t}{\Delta y\cdot\mu}\left[E_{z,i}\Big|^n_{i+1,j+1,k_a}-E_{z,i}\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}\right], \quad (16)$$

$$H_y\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}=\left\{H_y\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}\right\}_{FCC}-\frac{\Delta t}{\Delta x\cdot\mu}\left[E_{z,i}\Big|^n_{i+1,j+1,k_a}-E_{z,i}\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}\right]-\frac{\Delta t}{\Delta z\cdot\mu}\left[E_{x,i}\Big|^n_{i+1,j+1,k_a}+E_{x,i}\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}\right], \quad (17)$$

$$H_z\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}=\left\{H_z\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}\right\}_{FCC}-\frac{\Delta t}{\Delta y\cdot\mu}\left[E_{x,i}\Big|^n_{i+1,j+1,k_a}-E_{x,i}\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}\right]+\frac{\Delta t}{\Delta x\cdot\mu}\left[E_{y,i}\Big|^n_{i+1,j+1,k_a}-E_{y,i}\Big|^n_{i+\frac{1}{2},j+\frac{1}{2},k_a}\right], \quad (18)$$

where $\left\{H_x\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}\right\}_{FCC}$, $\left\{H_y\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}\right\}_{FCC}$ and $\left\{H_z\Big|^{n+1/2}_{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}}\right\}_{FCC}$ are the update equations of magnetic field components at location $\left\{i+\frac{3}{4},j+\frac{3}{4},k_a+\frac{1}{4}\right\}$ based upon FCC grids, respectively.

IV. NUMERICAL RESULTS

In this section, we will first validate update equations for the electric and magnetic field component reformulated on FCC grids, as the implementation of the proposed TF-SF boundary for the FDTD method are based on these update equations. Here, let us consider an empty rectangular metal cavity. The cavity size is $10mm\times15mm\times16mm$. The FCC cell sizes are $\delta=\Delta x=\Delta y=\Delta z=0.5mm$. The time step is $\Delta t=\delta/2.5c_0=0.667\,ps$, where c_0 is the speed of light in vacuum. A modulated Gaussian pulse is used as the source function, which is described by:

$$P(t)=\exp[-4\pi(t-t_0)^2/\tau^2]\times\cos(\omega t), \quad (19)$$

with the angular frequency parameter $\omega=4\pi\times10^{10}$, delay parameter $\tau=30\Delta t$ and $t_0=0.8\tau$. The excitation source is placed at (0,0,0). Figure 4 plots the waveform of E_z component of TM_{111} mode at observation point (0,0,10). The simulated results demonstrate the convergence of the proposed FCC grid method. In Fig. 5, the resonant frequency of TM_{111} mode of the cavity using the FCC method is 20.086GHz, and the theoretical resonant frequency of the cavity is 20.319GHz. The structure is also simulated in the standard Cartesian grids using the same cell sizes, and the resulting resonant frequency is 19.932GHz. The relative error is defined as $\delta_E=\left|(f_0-f_r)/f_0\right|\times100\%$, where f_0 is analytical frequency of the cavity, f_r is the simulation frequency of the cavity. The relative error obtained by FCC method and Cartesian method are 1.147% and 1.904%, respectively. It is shown that the FCC grid code is more accurate as compared to the traditional Yee grid method.

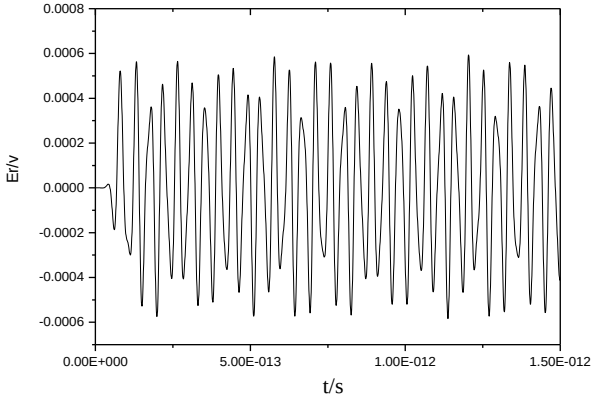


Fig. 4. E_z component of TM_{111} mode at observation point (0,0,10).

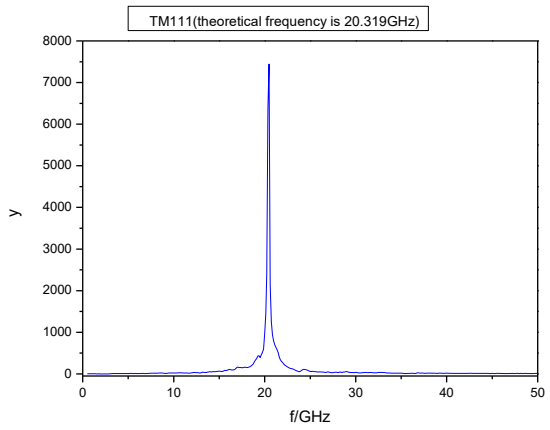


Fig. 5. The resonant frequency of TM_{111} mode of the cavity based upon the FCC grids.

Next, numerical examples are presented to validate the proposed TF-SF boundary for a plane wave excitation in the FDTD scheme based upon FCC grids. Here, a sinusoidal wave as an incident source is expressed by:

$$f = \sin(2\pi c_0 t / \lambda), \tag{20}$$

where c_0 is the speed of light in vacuum, and the wavelength of plane wave is λ=0.1m . The total computational region contains $280\times280\times280$ FCC cells with 280 cells along x, y and z directions, respectively. The TF region has 60 FCC cells along x, y and z directions, respectively. The spatial step is set as $\delta = \Delta x = \Delta y = \Delta z = 5\times10^{-3} m$, and the time step size is $\Delta t = \delta / 4c_0$=4.17 ps . The calculation program was set to run 1000 time steps. Numerical results in Fig. 6 are for the case of vertical incidence. Figures 6 (a) and 6 (b) show the plane wave amplitude distribution in the yz-plane section of the TF area at time step $220\Delta t$ and $380\Delta t$, respectively. The value of the right color bar indicates the relative amplitude value of E_x . The E_x field value against time at the observation point (0,0,0) is shown in Fig. 6 (c). It is clearly seen that there are no field leakage errors across the TF-SF boundary. The incident wave is successfully introduced into the TF area with the proposed TF-SF boundary.

Finally, the effectiveness of this TF-SF boundary for introducing plane wave at oblique incidence will be verified. The plane wave amplitude distribution in the yz-plane section of the TF area at time step $t = 220\Delta t$ and $t = 440\Delta t$ are shown in Fig. 7 (a) and Fig. 7 (b), respectively. Figure 7 (c) plots the E_x field value against time at the observation point (0,0,0). It is further shown that the incident wave is successfully introduced into the TF area efficiently with the proposed TF-SF boundary. It therefore appears that this study lays a solid foundation for using FDTD method based on FCC grids to deal with scattering problems. In Figs. 6 (c) and 7 (c), we found that the value of electric field varies slightly with the increase of time. This is due to the fact that the computational domain wasn't truncated by the absorption boundary. There are certain reflections from the computational boundary. The related absorbing boundary conditions for FDTD method based upon FCC grid will be studied in future work.

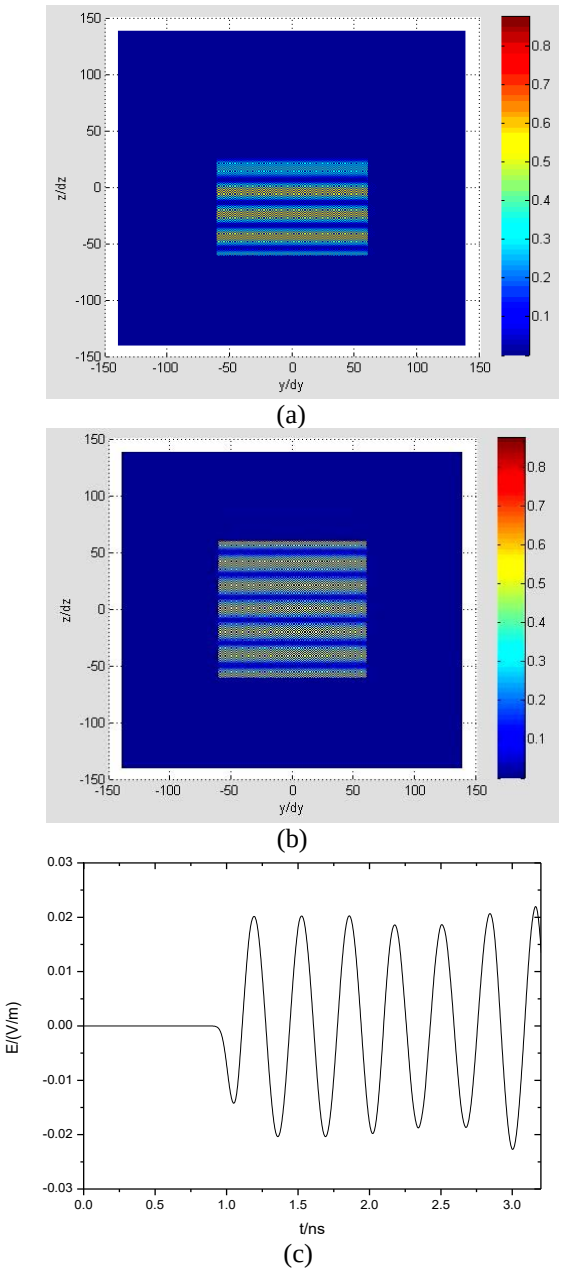


Fig. 6. Plane wave introduced into the TF area at vertical incidence: (a) amplitude distribution of electric component E_x at time step $t = 220\Delta t$; (b) amplitude distribution of electric component E_x at the time $t = 380\Delta t$; (c) field value of E_x against time at the observation point (0,0,0).

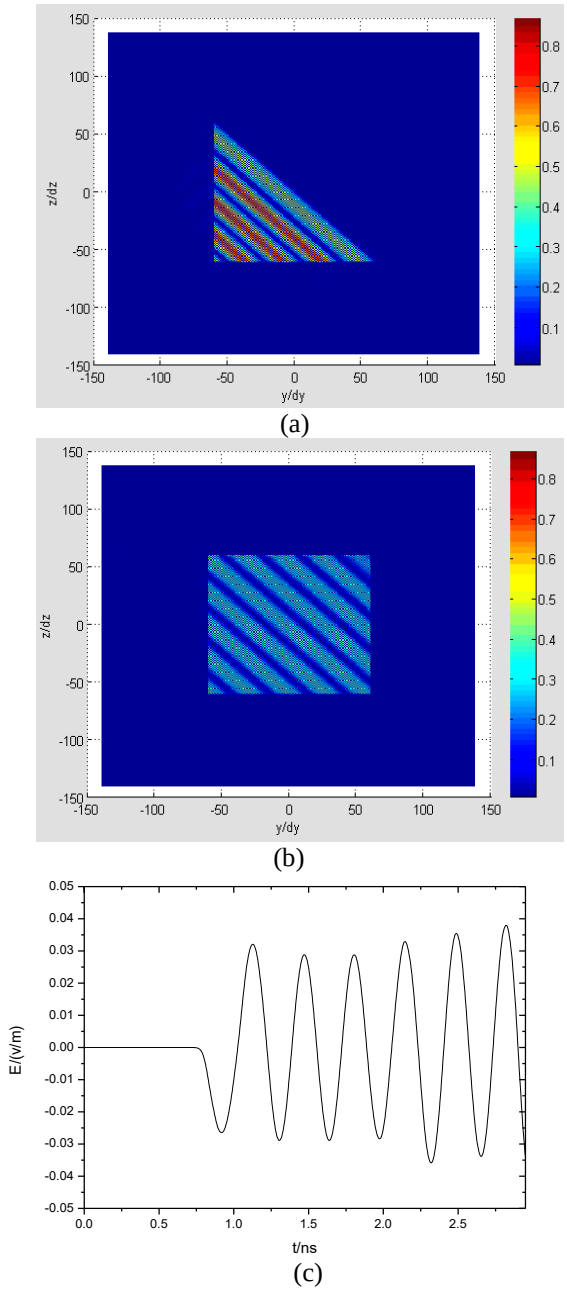


Fig. 7. Plane wave introduced into the TF area at vertical incidence: (a) amplitude distribution of electric component E_x at time step $t = 220\Delta t$; (b) amplitude distribution of electric component E_x at the time $t = 380\Delta t$; (c) field value of E_x against time at the observation point (0,0,0).

VI. CONCLUSION

In this paper, we formulate an effective TF-SF technique for the FDTD method based upon FCC grids is presented. According to the location of electric and magnetic field components in the FCC voxel, the difference iteration formulas for the electric and magnetic field components of the upper boundary are derived. Numerical results demonstrate that the incident plane wave can be successfully introduced into the TF area with the proposed TF-SF boundary. Future work will establish the complete FCC-FDTD algorithm for the EM scattering of targets by researching on the related absorbing boundary conditions and the near-to-far field transformation boundary.

ACKNOWLEDGMENT

This paper is supported by the National Natural Science Foundation of China (Grant No. 61601205, No. 11604126), Basic Research Program of Jiangsu Province (Natural Science Foundation for Young Scholars) (Grant No. BK20160541), and the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 16KJB140003).

REFERENCES

[1] Y. Q. Zhang and D. B. Ge, "A unified FDTD approach for electromagnetic analysis of dispersive objects," *Progress in Electromagnetics Research*, vol. 96, pp. 155-172, 2009.

[2] M. Wang, M. Yu, Z. Xu, et al., "Propagation properties of terahertz waves in a time-varying dusty plasma slab using FDTD," *IEEE Transactions on Plasma Science*, vol. 43, no. 12, pp. 4182-4186, 2015.

[3] A. Guellab and W. Qun, "High-order staggered finite difference time domain method for dispersive Debye medium," *ACES Journal*, vol. 33, no. 4, pp. 430-437, 2018.

[4] W. Shao, X. Ding, and B. Z. Wang, "Efficient sub-gridded FDTD for three-dimensional time-reversed electromagnetic field shaping," *ACES Journal*, vol. 33, no. 8, pp. 828-834, Aug. 2018.

[5] Z. Sun, L. Shi, Y. Zhou, B. Yang, and W. Jiang, "FDTD evaluation of LEMP considering the lossy dispersive ground," *ACES Journal*, vol. 33, no. 1, pp. 7-14, Jan. 2018.

[6] M. R. Cabello, L. D. Angulo, J. Alvarez, A. R. Bretones, and S. G. Garcia, "A new conformal FDTD for lossy thin panels," *IEEE Trans. Antennas Propag.*, vol. 67, no. 12, pp. 7433-7439, Dec. 2019.

[7] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-Domain Method*. 3rd ed., Artech House, Norwood, MA, 2005.

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

[9] M. E. Potter, M. Lamoureux, and M. D. Nauta, "An FDTD scheme on a face-centered-cubic (FCC) grids for the solution of the wave equation,"

Journal of Computational Physics, vol. 230, pp. 6169-6183, 2011.

[10] X. Fei, X. H. Tang, and X. J. Zhang, "The construction of low-dispersive FDTD on hexagon," *IEEE Trans. Antennas Propag.*, vol. 53, no. 11, pp. 3697-3703, 2005.

[11] Y. Liu, "Fourier analysis of numerical algorithms for the Maxwell equations," *Journal of Computational Physics*, vol. 124, pp. 396-416, 1996.

[12] M. E. Potter, and M. D. Nauta, "FDTD on face-centered cubic (FCC) grids for Maxwell's equations," *IEEE Trans. Antennas Progag.*, vol. 61, pp. 2116-2122, 2013.

[13] H. Kim, I. S. Koh, and J. G. Yook, "Enhanced total-field/scattered-field technique for isotropic-dispersion FDTD scheme," *IEEE Trans. Antennas Progag.*, vol. 58, no. 10, pp. 3407-3411, 2010.

[14] G. S. Singh, E. L. Tan, and Z. N. Chen, "Analytic fields with higher-order compensations for 3-D FDTD TF/SF formulation with application to beam excitations," *IEEE Trans. Antennas Propagat*., vol. 59, no. 7, pp. 2588-2598, 2011.

[15] T. P. Stefanski, "Application of the discrete Green's function-based antenna simulations for excitation of the total-field/scattered-field interface in the FDTD method," *Microwave and Optical Technology Letters*, vol. 56, no. 8, pp. 1949-1953, 2014.

[16] Y. Zhang, Y. M. Liu, and X. P. Li, "Total-field/scattered-field formulation for FDTD analysis of plane-wave propagation through cold magnetized plasma sheath," *IEEE Trans. Antennas Propag.*, vol. 68, no. 1, pp. 377-387, 2020.

Li Juan Shi received the M.S. degree in Physics from Xidian University, and the Ph.D. degree in Electrical and Mechanical Engineering from Jiangsu University, China, in 2006 and 2014, respectively. Her research interests include electromagnetic scattering and radiation, computational electromagnetism, and electromagnetic field theory.

Li Xia Yang received the Ph.D. in Electronics Engineering from Xidian University, China, in 2004. He is a Professor with the School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China. His research interests include numerical calculation of electromagnetic fields and antenna design.

Benchmark Electromagnetic Inverse Scattering by Using Differential Evolution – A Big Data Perspective

Anyong Qing

School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China
qinganyong@tsinghua.org.cn

Abstract **–** The benchmark electromagnetic inverse scattering problem is re-visited in this paper from a big data perspective. It serves as the benchmark application problem in systematic parametric study of differential evolution (DE). Representative strategies with a full sweeping of intrinsic control parameters are applied to draw a systematic picture of DE. Insights extracted from preliminary numerical results are presented to rebut the questionable statements and advise DE applicants.

Index Terms **–** Benchmark electromagnetic inverse scattering, differential evolution, parametric crime, parametric study, stochastic crime.

I. INTRODUCTION

A. Questionable statements in DE

Among the rapid expanding family of natural optimization algorithms, DE [1],[2] proposed by Price and Storn in 1995 is a very simple but very powerful evolutionary algorithm. It quickly earned its reputation as a prominent function optimizer through self-assessment and international showdowns. It has been applied to electromagnetic inverse scattering [3],[4], antennas [5], electromagnetic composite materials [6], frequency selective surfaces [12], microwave absorbers [8], and a lot of other mathematical and engineering optimization problems [9],[10],[11],[12],[13].

The standard notation DE/x/y/z of DE strategies [14] implies vast variants. As a matter of fact, only two operators, differential mutation and crossover, are involved in the standard notation that it is unable to cover all variants of DE strategies. Evolution mechanism, parental selection, and survival selection are missing.

There are two mechanisms to evolve the population in DE. The classic DE (CDE), also known as two-array method, applies static one while the dynamic DE (DDE) [15], or one-array method [16], evolves the population dynamically. A close analogy between the relationship between CDE and DDE and that between Jacobi and Gauss-Seidel method in linear algebra [17] can be made. Although it is well known that Gauss-Seidel method might converge faster more reliably than Jacobi method, it is claimed that there is "no dramatic difference in performance between the one- and two-array methods".

Differential mutation has been established as the crucial evolutionary operator leading to the success of DE. Its generic formulation to generate a mutant $\mathbf{v}^{n+1,i}$ for mother $\mathbf{p}^{n,i}$ reads:

$$\mathbf{v}^{n+1,i} = \mathbf{b}^{n,i} + \sum_{y\geq 1} F_y \left(\mathbf{d}^{n,i,y,1} - \mathbf{d}^{n,i,y,2} \right),$$

where $\mathbf{p}^{n,i} \in \mathbf{P}^n$, the population of the nth generation, $\mathbf{b}^{n,i}$ is the differential mutation base, $\mathbf{d}^{n,i,y,1}$ and $\mathbf{d}^{n,i,y,2}$ form the yth pair of donors, and F_y is the mutation intensity, also known as scale factor, for the yth vector difference. Please note that the notations of population individuals in this paper interchangeably represent the individuals as well as their N-dimensional vector of optimization parameters $\mathbf{x}$ whenever possible.

It has been a consensus in DE that different differential mutation bases balance the exploration and exploitation processes in DE differently. Accordingly, strategies with different differential mutation bases may have different performance. An anonymous reviewer comments that "in practice, DE/rand/1 is the most widely used strategy. Moreover, DE/best/1 is more prone to being trapped in a local optimum".

The successful innovation of differential mutation unfortunately shades other important ideas in DE as crossover does in genetic algorithms (GA). One of the victim operators in DE is crossover. It has been claimed that "The crossover method is not so important although Ken Price claims that binomial is never worse than exponential" [18].

The above highlighted statements have been well circulated in DE community. However, accumulating evidences pose stronger and stronger challenge against them. Serious measures have to be taken to examine these dubious statements to avoid potential damages to applicants' confidence in DE.

B. A big data practice

In 2004, after applying DE in electromagnetics for four years, this author started to get annoyed by the unsatisfactory performance of DE and inconsistent claims about DE strategies and intrinsic control parameters. An ambitious effort to reveal the relationship

Submitted On: December 25, 2016
Accepted On: June 9, 2017

between DE strategies, their intrinsic control parameters and mathematical features of optimization problems was triggered.

One of the fundamental activities in this effort is a literature survey on DE [9],[10],[13]. Mining and reviewing ever-growing literatures on DE is undoubtedly a big data process. It has been going on until today although there is a short break in 2012 due to this author's transfer from National University of Singapore to University of Electronic Science and Technology of China and a burglary into this author's office. Full details of the literature survey will be given in this author's coming monography.

Another essential activity in the effort is a systematic parametric study on DE [9],[10],[19],[20]. DE strategies, their corresponding intrinsic control parameters, termination conditions, toy functions and benchmark application problems, their corresponding non-intrinsic control parameters, form a testing system. Three performance indicators, reliability, efficiency, and robustness are defined to quantify performance of tested DE strategies.

This effort fits perfectly into the framework of big data. During this process, existing data is collected and huge amount of new data is generated. Most importantly, all available data is mined for insights to have more pleasant experience in future applying DE and develop better DE strategies.

C. Benchmark electromagnetic inverse scattering problem

Electromagnetic inverse scattering [21] are of great interest to both scientific researchers and engineers. Locating multiple two-dimensional perfectly conducting objects illuminated by TM-z plane waves and reconstructing their shape is a benchmark electromagnetic inverse scattering problem [4]. Scattered electric fields of some representative objects measured in controlled laboratory environment are also available [22],[23].

The benchmark electromagnetic inverse scattering problem has been solved by using a variety of inversion algorithms such as Newton-Kontorovitch algorithm [24], binary genetic algorithm [25], real-coded genetic algorithm [21], DE [4], and differential evolution with individuals in groups (GDE) [26]. Under the persistent promotion of this author, it has been accepted by both electromagnetic and optimization communities as one of the benchmark electromagnetic optimization problems due to its practical value and affordable computational cost.

D. Contributions of this paper

In this paper, the benchmark electromagnetic inverse scattering problem is re-visited. It serves as the benchmark application problem in systematic parametric study of DE. Representative strategies related with the aforementioned questionable statements are applied to solve the benchmark electromagnetic inverse scattering problem. Full sweeping of representative intrinsic control parameters has been conducted to draw a systematic picture of DE. Insights extracted from preliminary numerical results are presented to rebut the questionable statements and advise DE applicants. Moreover, parametric and stochastic crimes are defined to promote appropriate practice of applying and comparing stochastic and/or intrinsic control parameters-dependent optimization algorithms.

II. BENCHMARK ELECTROMAGNETIC INVERSE SCATTERING PROBLEM

A. Configuration

For better readability and completeness of this paper, the problem geometry is re-depicted in Fig. 1, where O is the origin of the global coordinate system, Ω, a circle of radius R^{meas}, is the measuring (data) domain in which the scattered electric fields are measured, the black dots on Ω are receivers, D is the imaging (object) domain which is usually chosen to be circular or rectangular.

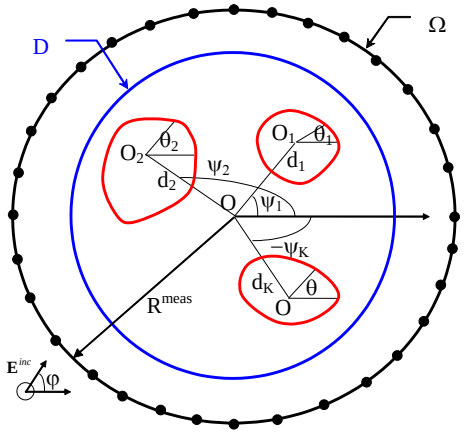


Fig. 1. Geometry of the electromagnetic inverse scattering problem.

The objects of interest are the K perfectly conducting cylinders in D, $O_i(d_i, \psi_i)$ is the local origin of the ith cylinder which can be any point within the cylinder contour C_i, d_i is the distance between O and O_i, ψ_i is the angle from horizontal or $+x$ direction to vector $\overrightarrow{OO_i}$. C_i is represented by a local shape function $\rho_i = F_i(\theta_i)$ in the local polar coordinate system. Physically, $F_i(\theta_i) \geq 0\ \forall \theta_i \in [0,2\pi]$. Apparently, the same contour can be represented by infinite sets (O_i, F_i).

The local shape function $F_i(\theta_i)$ used to be approximated by a trigonometric series $F_i^T(\theta_i)$ of order $N/2$ [24],[25]:

$$F_i^T(\theta_i) = \sum_{j=0}^{N/2} A_{ij} cos(j\theta_i) + \sum_{j=1}^{N/2} A_{i(j+N/2)} sin(j\theta_i),$$

where A_{ij} are coefficients.

Five contours are generated as shown in Fig. 2 by setting the coefficients as a random number uniform in the suggested ranges [25]. The local shape functions are shown in the left figure while the contours are shown in the right one. Only the local shape function in red dash line is non-negative. The corresponding contour in red color is regular. Apparently, it is very hard for $F_i^T(\theta_i)$ to meet the physical requirement on local shape functions.

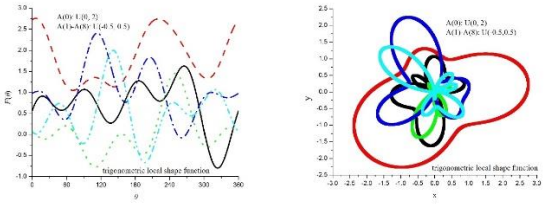


Fig. 2. Trigonometric local shape functions.

In 2000, Qing [27],[28] proposed the closed cubic B-splines local shape functions $F_i^B(\theta_i)$ with N control points C_{ij} to approximate $F_i(\theta_i)$:

$$F_i(\theta_i) \approx F_i^B(\theta_i) = \sum_{j=0}^{3} C_{i\,\mathrm{mod}(m-1+j,N)} Q_i(t),$$

where

$$Q_0(t) = \frac{1}{6}(1-t)^3,$$
$$Q_1(t) = \frac{1}{2}t^3 - t^2 + \frac{2}{3},$$
$$Q_2(t) = -\frac{1}{2}t^3 + \frac{1}{2}t^2 + \frac{1}{2}t + \frac{1}{6},$$
$$Q_3(t) = \frac{1}{6}t^3,$$
$$t = \frac{N}{2\pi}\theta_i - m,$$
$$m = \left\lfloor \frac{N}{2\pi}\theta_i \right\rfloor.$$

It has been proven that:

$$\min_{0\le j\le N-1} C_{ij} \le F_i^B(\theta_i) \le \max_{0\le j\le N-1} C_{ij}.$$

Therefore, by setting $\min_{0\le j\le N-1} C_{ij} \ge 0$, non-negative definiteness of $F_i^B(\theta_i)$ can be guaranteed.

Similarly, five contours are generated as shown in Fig. 3 by setting the control points as a random number uniform in the specified ranges. The local shape functions are shown in the left figure. All of them are non-negative. In addition, the contours are shown in the right one. All contours are regular.

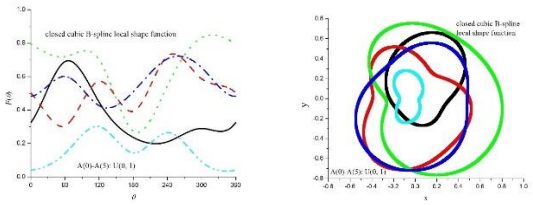


Fig. 3. Closed cubic B-spline local shape functions.

The objects of interest are illuminated by TM plane waves (time factor $e^{j\omega t}$ assumed and suppressed where $\omega = 2\pi f$ is the angular frequency) of unit amplitude:

$$\mathbf{E}^i(\mathbf{r}) = \hat{z} exp\left(-jk_0 \hat{k}\cdot\mathbf{r}\right),$$

where $\mathbf{r} = x\hat{x} + y\hat{y}$, $k_0 = \omega/c$ is the wave number in free space, c is the light speed, $\hat{k} = cos\varphi\hat{x} + sin\varphi\hat{y}$ is the incident wave unit vector and φ is the incident angle. $\hat{x}$, $\hat{y}$ and $\hat{z}$ are the unit vectors in the x, y and z directions respectively.

B. Direct problem

The electric field integral equations governing the scattering problem are:

$$E^i(r) = \sum_{j=1}^{K} \frac{\omega\mu_0}{4} \iint_{C_j} J_j(r') H_0^{(2)}\left(k_0|r-r'|\right) dr', r \in \bigcup_{i=1}^{K} C_i,$$

$$E^s(r) = \sum_{j=1}^{K} -\frac{\omega\mu_0}{4} \iint_{C_j} J_j(r') H_0^{(2)}\left(k_0|r-r'|\right) dr', r \in \Omega,$$

where μ_0 is the permeability of free space, $J_j(\mathbf{r}')$ is the induced surface current intensity on the surface of the jth cylinder, $H_0^{(2)}(\cdot)$ is the second kind Hankel's function of zeroth order.

C. Inverse problem

The inverse problem is to locate the objects of interest and reconstruct their shape, given the measured scattered electric fields $\mathbf{E}^{sm}$ which is an $N_f \times N_a \times N_r$-dimensional vector $\left[E_1^{sm}, \cdots, E_j^{sm}, \cdots, E_{N_f\times N_a\times N_r}^{sm}\right]$, N_f, N_a, and N_r are the total number of frequencies, incident angles and receivers respectively. It is cast into an unconstrained functional minimization problem whose optimization parameters are:

$$\mathbf{x} = \bigcup_{i=1}^{K}\left[d_i, \psi_i, F_i(\theta_i)\right] = \left[\mathbf{x}_1, \cdots, \mathbf{x}_K\right],$$

and objective function is:

$$f(\mathbf{x}) = \frac{\left\|\mathbf{E}^{sm} - \mathbf{E}^s(\mathbf{x})\right\|}{\left\|\mathbf{E}^{sm}\right\|},$$

where $\mathbf{x}_i = [d_i, \psi_i, F_i(\theta_i)]$, $\mathbf{E}^s(\mathbf{x})$ is an $N_f \times N_a \times N_r$-dimensional vector of scattered fields corresponding to the profile $\cup_{i=1}^{K}[d_i, \psi_i, F_i(\theta_i)]$,

$$\left\|\mathbf{E}^{sm}\right\| = \sqrt{\sum_{j=1}^{N_f \times N_a \times N_r} \left|E_j^{sm}\right|^2},$$

$$\left\|\mathbf{E}^{sm} - \mathbf{E}^{s}\right\| = \sqrt{\sum_{j=1}^{N_f \times N_a \times N_r} \left|E_j^{sm} - E_j^{s}\right|^2}.$$

One of the distinctive mathematical features of functional $f(\mathbf{x})$ lies with its non-uniqueness, i.e., there are infinite minima corresponding to the infinite sets of (O_i, F_i) representing the same contours C_i. The electromagnetic equivalence principle makes the non-uniqueness issue more complicated.

Approximating $F_i(\theta_i)$ by $F_i^B(\theta_i)$ simplifies the unconstrained functional minimization problem into an unconstrained parameter minimization problem whose objective function is intact but the optimization parameters are:

$$\mathbf{x} = \left[\mathbf{x}_1, \cdots, \mathbf{x}_K\right],$$

where $\mathbf{x}_i = \left[d_i, \psi_i, C_{i0}, \cdots, C_{i(N-1)}\right]$.

III. DIFFERENTIAL EVOLUTION

A. General framework

As seen from Fig. 4, CDE and DDE share initialization, differential mutation with parental selection embedded, crossover, evaluation, and survival selection, but differ in evolution. Initial population $\mathbf{P}^0$ is generated through initialization. In CDE, population evolves generation by generation until at least one of the termination conditions is met, hopefully the objective is met. On the other hand, although index n is still in use in DDE for better clarity, it does not bear the same meaning as generation index in CDE. More importantly, the population continuously updates itself.

<table>
<tr><td>initialization</td><td colspan="2">$n = 0$
do $i = 1, N_p$
do $j = 1, N$
$x_j^{0,i} = a_j + r_j^{0,i}\left(b_j - a_j\right)$
end do
end do</td></tr>
<tr><td rowspan="2">evolution</td><td>CDE</td><td>DDE</td></tr>
<tr><td>do while (termination conditions not satisfied)
$n = n + 1$
do $i = 1, N_p$
differential evolution to get $\mathbf{v}^{n+1,\,i}$
crossover to get $\mathbf{c}^{n+1,\,i}$
evaluation of $\mathbf{c}^{n+1,\,i}$
survival selection to get $\mathbf{p}^{n+1\,i}$
end do
end do</td><td>do while (termination conditions not satisfied)
$i = 1$
differential evolution to get $\mathbf{v}^{n+1,\,i}$
crossover to get $\mathbf{c}^{n+1,\,i}$
evaluation of $\mathbf{c}^{n+1,\,i}$
survival selection to get $\mathbf{p}^{n+1\,i}$
if (i .eq. N_p) then
$n = n + 1$
$i = 1$
else
$i = i + 1$
end if
end do</td></tr>
</table>

Fig. 4. Pseudocode of differential evolution.

B. System of parametric study

1) Differential evolution strategies

Strategies in the system are classified into four categories according to their evolution and learning mechanism: CDE, DDE, CDE with opposition-based learning (OCDE) and DDE with opposition-based learning (ODDE) [29],[30]. An expanded notation system, i/x/(y,d)/z/s, is implemented to represent the specific operations in each strategy where x, y, and z bear the same meaning as in standard DE notation, i stands for initialization, d stands for donor selection, and s stands for survival selection. Therefore, the concerned DE strategies are:

(a) CDE/i/x/(y,d)/z/s,
(b) DDE/i/x/(y,d)/z/s,
(c) OCDE/i/x/(y,d)/z/s,
(d) ODDE/i/x/(y,d)/z/s.

It has to be pointed out that even the above expanded notation system cannot cover all DE strategies reported in literatures.

2) Intrinsic control parameters

DE is population-based. Therefore, all DE strategies share the same intrinsic control parameter, population size N_p. Two sets of population sizes, {8, 16, 24, 32, 40, 48, 56, 64, 72, 80, 120, 160, 200, 400, 800, 1600} and

{10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, 1000, 2000} are implemented in the parametric study according to the dimension of the test problem. The first set is applied for the benchmark electromagnetic inverse scattering problem re-visited in this paper.

It is well accepted that differential mutation is one of the key innovations leading to the success of DE. From this point of view, differential mutation has to be present in all DE strategies. In accordance, all DE strategies share at least one more intrinsic control parameter, mutation intensity F, or mutation intensities F_y if there are more than one pair of donors. The tested cases of mutation intensity form a set as $\{F|F = j\Delta_m, 1 \leq j \leq (F^U - F^L)/\Delta_m\}$ where F^L and F^U are the lower and upper bounds. The step size Δ_m is usually adjusted according to the computational cost of the test problem. $F^L = 0$, $F^U = 1$ and $\Delta_m = 0.1$ are applied for the benchmark electromagnetic inverse scattering problem re-visited in this paper.

Traditionally, crossover is present in almost all applications of DE. In this regard, crossover probability p_c, the companion intrinsic control parameter to crossover, is also essential and shared. Similarly, the tested cases of crossover probability form a set as $\{p_c|p_c = j\Delta_c, 1 \leq j \leq (p_c^U - p_c^L)/\Delta_c\}$ where $0 \leq p_c^L < p_c^U \leq 1$. The step size Δ_c is also adjusted according to the computational cost of the test problem, usually in synchrony with Δ_m. $p_c^L = 0$, $p_c^U = 1$ and $\Delta_c = 0.1$ are applied for the benchmark electromagnetic inverse scattering problem re-visited in this paper.

Obviously, a specific DE strategy may have more companion intrinsic control parameters.

3) Test problems

Test problems in the system include toy functions and benchmark application problems. A literature survey on test problems was started at almost the same time as the literature survey on DE. As of Dec. 22, 2016, more than 500 toy functions and 100 application problems from different disciplines, including the benchmark electromagnetic inverse scattering problem re-visited in this paper, have been collected.

From the point of view of parametric study, landscape and mathematical features of test problems and their relationship with other components in the system for parametric study are more fundamental. Features such as decomposability, modality, continuity, differentiability, dimensionality, uniqueness, and many more are under careful scrutiny.

One of the control parameters companion to test problems is the search space of their optimization parameters. Whenever possible, a wider search space is more welcome because it imposes less requirement on *a priori* knowledge, especially for practical engineering optimization problems. The standard setting of $0 < d_i \leq 1\lambda$, $0 \leq \psi_i \leq 2\pi$, and $0 \leq C_{ij} \leq 1\lambda$, for the benchmark electromagnetic inverse scattering problem, is implemented in this paper.

4) Termination conditions

When optimum (minimum by default) of the optimization problem under test is known, it is very natural to terminate the search when the objective is met. It is mathematically formulated as:

$$f(\mathbf{x}) - f(\mathbf{x}^*) \leq \varepsilon,$$

where $f(\mathbf{x}^*)$ is the known minimum, and ε is the threshold value to reach. For most engineering optimization problem, $\varepsilon = 0.01$.

The optimum of some toy functions and most benchmark application problems is yet to determine. In this case, the second termination condition, time limit, is introduced straightforwardly. Due to its difficulty to implement across platforms, limit of number of objective function evaluations has been proposed as an equivalent alternate. In this paper, it is set as 2000 times the problem dimension.

5) Performance indicators

Three performance indicators, reliability, efficiency, and robustness are defined in descending order of priority to quantify performance of tested DE strategies. Reliability refers the capability to find the optimum of the concerned optimization problem, efficiency refers to the number of objective function evaluations before the optimum is successfully located, while robustness refers to the sensitivity of reliability and efficiency with respect to intrinsic control parameters of concerned optimization algorithms and/or other control parameters in the parametric study system. Due to page limitation, explicit and physically meaningful quantitative definition of the three performance indicators will be presented in a new book by this author, hopefully published by John Wiley & Sons and/or IEEE Press.

IV. NUMERICAL RESULTS AND INSIGHTS

A. Profile reconstruction

All profiles considered in [4] were re-simulated here. Due to non-uniqueness of the problem and stochastic nature of DE, special attention is given to consistency of successfully reconstructed profiles to address the effect of non-uniqueness and randomness.

Although the final optimization parameters **x** and the corresponding objective function value $f(\mathbf{x})$ obtained by successful searches are diverse, in terms of the final reconstructed profiles, all participating DE strategies perform perfectly. In all noiseless synthetic reconstruction, true profiles are successfully reconstructed to the acceptable engineering accuracy ε if DE converges. In real reconstruction, the objective function value cannot go below ε because of noise in the measured scattered

electric fields. As a matter of fact, $f(\mathbf{x}^*) = h > \varepsilon$ where $\mathbf{x}^*$ is the true profile. Fortunately, the reconstructed profiles $\mathbf{x}$ agree quite well with those given in [4] when $f(\mathbf{x}) \le h + \varepsilon$.

B. Parametric study

Our attention in this paper is focused on evaluating DE instead of solving the benchmark electromagnetic inverse scattering problem. Therefore, besides the reconstructed profiles, we are more interested in the statistical characteristics of the reconstruction process.

The representative reliability and efficiency of DE is shown in Fig. 5. For consistency with [9], reliability here is represented by the percentage of total number of successful searches (TNSS) among all searches and the percentage of number of successful trials (NST) among all trials. Similarly, efficiency in Fig. 5 is defined as the minimal average number of objective function evaluations (MANOFE) of all successful trials. The presented results show DDE significantly outperforms CDE in terms of both reliability and efficiency.

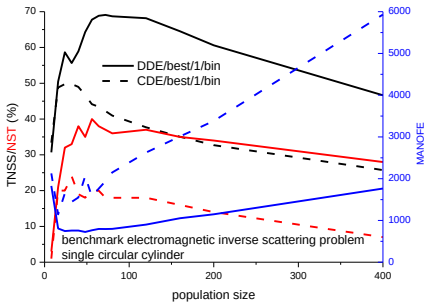


Fig. 5. Reliability and efficiency of differential evolution solving benchmark electromagnetic inverse scattering of a perfectly conducting circular cylinder.

It is also observed from obtained numerical results that:

(a) DE strategies with best differential mutation base show better performance than those with random differential mutation base;

(b) DE strategies with exponential crossover perform better than those with binomial crossover. The aforementioned statements about crossover are inconsistent with numerical results.

Due to page limitation, full set of performance has been scheduled to be published in this author's next monograph.

1) Optimal intrinsic control parameters

It is observed from Fig. 5 that the reliability of DE presents a shape of bell with respect to population size. Thus, although a larger population may be more diverse, extra-large population does not necessarily mean better performance. 20-80 seems acceptable for DDE while the optimal population size of CDE belongs to a much narrower range due to its much stronger sensitivity of efficiency with respect to population size. In another word, DDE is more robust.

The optimal mutation intensity and crossover probability corresponding to the most efficient successful trial is shown in Fig. 6. DDE prefers remarkably weaker mutation which is reasonable because dynamic evolution introduces more diversity into the population. In addition, both CDE and DDE prefer full or nearly full crossover which implies an important role of crossover in DE. The stronger and more inconsistent change of optimal mutation intensity and crossover probability of CDE in the vicinity of its optimal population size further demonstrates DDE's robustness.

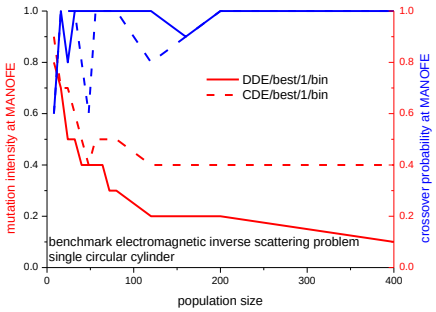


Fig. 6. Optimal mutation intensity and crossover probability.

2) Parametric crime

Robustness, or sensitivity of performance of DE with respect to its intrinsic control parameters at the vicinity of its optimal values has long been known an important issue. Intuitively, DDE can be observed from Fig. 5 more robust than CDE.

Essentially, the parametric study here involves two important aspects: identifying optimal values of intrinsic control parameter and quantifying robustness of DE at its vicinity. By now, this author is still working on a quantitative performance indicator to represent robustness. A promising candidate might be the safe zone borrowed from the idea of beam width or bandwidth in electromagnetics.

Although experiences show that DE is more robust, or less sensitive with intrinsic control parameters, underestimating or even ignoring it is by no means wise and might result in false and misleading conclusions. It implies flat performance throughout viable range of intrinsic control parameters which is absolutely not true. Such a misconduct is termed as parametric crime as sketched in Fig. 7 in analogy to inverse crimes in inverse problems.

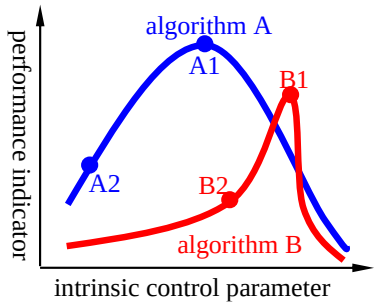


Fig. 7. Parametric crime.

Assume that algorithms A and B are compared with each other. Both of them involve an intrinsic control parameter. Obviously, algorithm A is better than B. However, if algorithm B at B1 is compared with algorithm A at A2, a false conclusion that algorithm B outperforms algorithm A will be drawn.

In Fig. 5, the efficiency of CDE for $N_p < 100$ is observed better than that of DDE for $N_p = 400$. Obviously, it is not true that CDE is better than DDE.

3) Stochastic crime

As a member of stochastic optimization algorithms, the process of a single run of DE is hardly reproducible. One might converge while the other fails although all settings are exactly the same. Likewise, two successful runs are very likely to have completely different converging process. In this regard, it would be a serious misleading conduct to define the performance of any stochastic optimization algorithm as the one of a single run. Similarly, such a misconduct is termed as stochastic crime.

Stochastic crime might also lead to reversed false conclusion about competing optimization algorithms. For example, in our practice, CDE only takes 720 objective function evaluations to solve the benchmark electromagnetic inverse scattering problem of a single perfectly conducting circular cylinder in the most efficient run, while one of the successful DDE runs takes 1280 objective function evaluations to solve the same problem. One might accordingly claim false advantage of CDE against DDE. In fact, the average number of objective function evaluations of CDE and DDE is 1248.8 and 625.07 respectively. The claim of advantage of CDE against DDE is apparently false.

V. CONCLUSIONS

In this paper, the benchmark electromagnetic inverse scattering problem is re-visited from a big data perspective. It serves as the benchmark application problem in systematic parametric study of DE. Preliminary numerical results re-confirm the advantage of dynamic evolution. Best differential mutation base and exponential crossover are also observed more competitive.

ACKNOWLEDGEMENT

The author wishes to thank financial and facility support from Singapore Defense Science and Technology Agency, Defense Research and Technology Office, Singapore, National Young Thousand Talent Programme of China, and University of Electronic Science and Technology of China.

This work was supported in part by the National Key R&D Plan under Grant 2018YFC0809500, National Young Thousand Talent Grants A0920502051826, YH199911041801, & YX1199912371901, Double First Class Grant YX1300112601801-5, and Foreign Talent in Culture and Education Grant 110000207520190055.

The author would also like to thank numerous colleagues, students, and collaborators helping the involved work in China, Singapore, Germany, France, UK, USA, and Canada.

REFERENCES

[1] R. Storn and K. V. Price, *Differential Evolution - A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces*, TR-95-012, International Computer Science Institute, Mar. 1995.

[2] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, Series: *Natural Computing Series*, Berlin: Springer, 2005.

[3] K. A. Michalski, "Electromagnetic imaging of circular-cylindrical conductors and tunnels using a differential evolution algorithm," *Microwave and Optical Technology Letters*, vol. 27, no. 5, pp. 330-334, Dec. 5, 2000.

[4] A. Qing, "Electromagnetic inverse scattering of multiple two-dimensional perfectly conducting objects by the differential evolution strategy," *IEEE Transactions on Antennas and Propagation*, vol. 51, no. 6, pp. 1251-1262, June 2003.

[5] S. Yang, Y. B. Gan, and A. Qing, "Sideband suppression in time modulated linear arrays by the differential evolution algorithm," *IEEE Antennas and Wireless Propagation Letters*, vol. 1, no. 9, pp. 173-175, 2002.

[6] A. Qing, X. Xu, and Y. B. Gan, "Anisotropy of composite materials with inclusion with orientation preference," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 2, pp. 737-744, Feb. 2005.

[7] X. F. Luo, A. Qing, and C. K. Lee, "Application of the differential-evolution strategy to the design of frequency-selective surfaces," *International Journal of RF and Microwave Computer-aided Engineering*, vol. 15, no. 2, pp. 173-180, 2005.

[8] A. Qing, "Design of thin wideband planar absorber

using dynamic differential evolution and real electromagnetic composite materials," *2011 IEEE AP-S International Symposium*, Spokane, WA, pp. 2912-2915, July 3-8, 2011.

[9] A. Qing, *Differential Evolution: Fundamentals and Applications in Electrical Engineering*, New York: John Wiley & Sons-IEEE Press, 2009.

[10] A. Qing and C. K. Lee, Eds., *Differential Evolution in Electromagnetics*, Berlin: Springer, 2010.

[11] S. Das and P. N. Suganthan, "Differential evolution: A survey of the state-of-the-art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4-31, Feb. 2011.

[12] P. Rocca, G. Oliveri, and A. Massa, "Differential evolution as applied to electromagnetics," *IEEE Antennas and Propagation Magazine*, vol. 53, no. 1, pp. 38-49, Feb. 2011.

[13] A. Qing, Comment on "P. Rocca, G. Oliveri, and A. Massa, "Differential evolution as applied to electromagnetics," *IEEE Antennas and Propagation Magazine*, vol. 53, no. 4, pp. 169-171, Aug. 2011.

[14] R. Storn and K. Price, "Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341-359, 1997.

[15] A. Qing, "Dynamic differential evolution strategy and applications in electromagnetic inverse scattering problems," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 1, pp. 116-125, Jan. 2006.

[16] D. Corne, M. Dorigo, and F. Glover, Eds., *New Ideas in Optimization*. London: The McGraw-Hill Companies, pp. 79-108, 1999.

[17] G. H. Golub and C. F. Van Loan, *Matrix Computations*. 3rd Ed., The Johns Hopkins University Press, 1996.

[18] http://www1.icsi.berkeley.edu/~storn/code.html, last accessed on Dec. 25, 2016.

[19] A. Qing, "A parametric study on differential evolution based on benchmark electromagnetic inverse scattering problem," *2007 IEEE Congress on Evolutionary Computation*, Singapore, pp. 1904-1909, Sep. 25-28, 2007.

[20] A. Qing, "A study on base vector for differential evolution," *2008 IEEE World Congress Computational Intelligence/2008 IEEE Congress Evolutionary Computation*, Hong Kong, pp. 550-556, June 1-6, 2008.

[21] A. Qing, C. K. Lee, and L. Jen, "Electromagnetic inverse scattering of two-dimensional perfectly conducting objects by real-coded genetic algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 665-676, Mar. 2001.

[22] K. Belkebir and M. Saillard, "Special section: testing inversion algorithms against experimental data," *Inverse Problems*, vol. 17, no. 6, pp. 1565-1571, 2001.

[23] J. M. Geffrin, P. Sabouroux, and C. Eyraud, "Free space experimental scattering database continuation: Experimental set-up and measurement precision," *Inverse Problems*, vol. 21, no. 6, pp. S117-S130, 2005.

[24] C. C. Chiu and Y. W. Kiang, "Microwave imaging of multiple conducting cylinders," *IEEE Transactions on Antennas and Propagation*, vol. 40, no. 8, pp. 933-941, Aug. 1992.

[25] C. C. Chiu and P. T. Liu, "Image reconstruction of a perfectly conducting cylinder by the genetic algorithm," *IEE Proceedings H-Microwaves, Antennas & Propagation*, vol. 143, no. 3, pp. 249-253, June 1996.

[26] A. Qing, "Electromagnetic inverse scattering of multiple perfectly conducting cylinders by differential evolution strategy with individuals in groups (GDES)," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 5, pp. 1223-1229, May 2004.

[27] A. Qing and C. K. Lee, "Application of cubic B-spline function in microwave imaging of a perfectly conducting cylinder by genetic algorithm," *2000 Workshop Microwave Imaging: Algorithms Applications, European Microwave Conference*, Paris, France, Oct. 2-6, 2000.

[28] A. Qing, "Microwave imaging of parallel perfectly conducting cylinders with transverse electric scattering data," *Journal of Electromagnetic Waves and Applications*, vol. 15, no. 5, pp. 665-685, 2001.

[29] H. R. Tizhoosh, "Opposition-based learning: A new scheme for machine intelligence," *2005 International Conference on Computational Intelligence for Modelling, Control and Automation/ 2005 International Conference on Intelligent Agents, Web Technologies, and Internet Commerce*, Vienna, Austria, vol. 1, pp. 695-671, Nov. 28-30, 2005.

[30] S. Rahnamayan, H. P. Tizhoosh, and M. M. A. Salama, "Opposition-based differential evolution," *2006 IEEE Congress on Evolutionary Computation, Vancouver*, Canada, pp. 2010-2017, 2006.

Application of PS-ADI-MRTD Method in Frequency Selective Surface Analysis

Yawen Liu and Pin Zhang

PLA University of Army Engineering, Nanjing, 210007, China
liuyawen1111@163.com, pinzhangthree@sina.com

Abstract ─ In this paper, a novel periodic spectral alternating direction implicit multi-resolution time domain (PS-ADI-MRTD) method is proposed for solving periodic structures. The algorithm for inversion of the block periodic tri-diagonal matrices is presented. A typical example is given to validate the effectiveness of the proposed algorithm, and the numerical results also show that the proposed algorithm can save more computation time than the traditional MRTD method. Furthermore, the PS-ADI-MRTD method is applied to analyze the electromagnetic scattering characteristics of two types frequency selective surface (FSS) structures.

Index Terms ─ Alternating Direction Implicit (ADI), Frequency Selective Surface (FSS), Multi-Resolution Time Domain (MRTD), spectral technique.

I. INTRODUCTION

The Multi-Resolution Time-Domain (MRTD) technique was first proposed by Krumpholz and Katehi [1-2]. Although with a high linear dispersion performance [3-5], the MRTD method has a major disadvantage that the time stability condition is stricter than FDTD [2], which limits the computational efficiency of the MRTD method. In 2001, Chen [6] proposed the ADI-MRTD method, which is unconditionally stable, that is, the choice of time step is not limited to the size of space interval.

In computational electromagnetics, there are various periodic structures, such as gratings, photonic bandgaps (PBG), frequency selective surfaces (FSS) and phased antenna arrays. And notable progress on the engineering applications of periodic structures has been achieved at the same time [7-13]. In this paper, the spectral technique [11-13] is applied to ADI-MRTD method, resulting in the periodic spectrum ADI-MRTD (PS-ADI-MRTD) algorithm. With the advantages of highly-linear dispersion performance and the CTW wave, it is suitable for calculating the periodic structures with high efficiency and oblique incidence. Moreover, a block periodic tridiagonal matrix inversion algorithm is presented. The scattering analysis of square thin plate arrays verifies the effectiveness and effectiveness of this method, and the numerical simulation shows that PS-ADI-MRTD has a good performance in saving computing time.

At the end, the PS-ADI-MRTD method is used to analyze the frequency scattering characteristics of a patch-type frequency selective surface and an aperture-type frequency selective surface, and the influence of incident angle and array arrangement on their dispersion characteristics is also discussed.

II. THE PROPOSED PS-ADI-MRTD ALGORITHM

The basic equations form of the PS-ADI-MRTD is similar to that of the ADI-MRTD [6]. In the PS-ADI-MRTD method, the Maxwell's equations are divided into two sub-steps, $(n+1/2)$ th and $(n+1)$ th step. Taking E_x and H_z as example, in the $(n+1/2)$ th sub-step, the following difference updating equations can be obtained as:

$$E_x^{\,n+\frac{1}{2}}\left(i+\tfrac{1}{2},j,k\right)=CA_x\left(i+\tfrac{1}{2},j,k\right)E_x^{\,n}\left(i+\tfrac{1}{2},j,k\right)+CB_x\left(i+\tfrac{1}{2},j,k\right)\left(\begin{array}{l}\dfrac{1}{\Delta y}\displaystyle\sum_l a(l)H_z^{\,n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+l+\tfrac{1}{2},k\right)\\ -\dfrac{1}{\Delta z}\displaystyle\sum_l a(l)H_y^{\,n}\left(i+\tfrac{1}{2},j,k+l+\tfrac{1}{2}\right)\end{array}\right), \tag{1}$$

$$H_z^{\,n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+\tfrac{1}{2},k\right)=CP_z\left(i+\tfrac{1}{2},j+\tfrac{1}{2},k\right)H_z^{\,n}\left(i+\tfrac{1}{2},j+\tfrac{1}{2},k\right)+CQ_z\left(i+\tfrac{1}{2},j+\tfrac{1}{2},k\right)\left(\begin{array}{l}\dfrac{1}{\Delta y}\displaystyle\sum_l a(l)E_x^{\,n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+l+1,k\right)\\ -\dfrac{1}{\Delta x}\displaystyle\sum_l a(l)E_y^{\,n}\left(i+l+1,j+\tfrac{1}{2},k\right)\end{array}\right). \tag{2}$$

In the $(n+1)$ th sub-step, the difference updating equations can be obtained as:

Submitted On: April 14, 2019
Accepted On: April 12, 2020

$$E_x^{n+1}\left(i+\tfrac{1}{2},j,k\right)=CA_x\left(i+\tfrac{1}{2},j,k\right)E_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j,k\right)$$
$$+CB_x\left(i+\tfrac{1}{2},j,k\right)\left(\begin{array}{l}\dfrac{1}{\Delta y}\sum_l a(l)H_z^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+l+\tfrac{1}{2},k\right)\\ -\dfrac{1}{\Delta z}\sum_l a(l)H_y^{n+1}\left(i+\tfrac{1}{2},j,k+l+\tfrac{1}{2}\right)\end{array}\right), \tag{3}$$

$$H_z^{n+1}\left(i+\tfrac{1}{2},j+\tfrac{1}{2},k\right)$$
$$=CP_z\left(i+\tfrac{1}{2},j+\tfrac{1}{2},k\right)H_z^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+\tfrac{1}{2},k\right)$$
$$+CQ_z\left(i+\tfrac{1}{2},j+\tfrac{1}{2},k\right)\left(\begin{array}{l}\dfrac{1}{\Delta y}\sum_l a(l)E_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+l+1,k\right)\\ -\dfrac{1}{\Delta x}\sum_l a(l)E_y^{n+1}\left(i+l+1,j+\tfrac{1}{2},k\right)\end{array}\right), \tag{4}$$

where

$$\left.\begin{array}{l}CA_x\left(i+\dfrac{1}{2},j,k\right)=\dfrac{4\varepsilon-\sigma\Delta t}{4\varepsilon+\sigma\Delta t}\\ CB_x\left(i+\dfrac{1}{2},j,k\right)=\dfrac{2\Delta t}{4\varepsilon+\sigma\Delta t}\\ CP_z\left(i+\dfrac{1}{2},j+\dfrac{1}{2},k\right)=\dfrac{4\mu-\sigma_m\Delta t}{4\mu+\sigma_m\Delta t}\\ CQ_z\left(i+\dfrac{1}{2},j+\dfrac{1}{2},k\right)=\dfrac{2\Delta t}{4\mu+\sigma_m\Delta t}\end{array}\right\}. \tag{5}$$

The coefficients $a(l)$ for $0\le l\le 2$ have been tabulated in [3], and the coefficients $a(l)$ for $l<0$ are given by the symmetry relation $a(-1-l)=-a(l)$:

$$E_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j,k\right)-\frac{CB_x}{\Delta y^2}$$
$$\times\sum_l a(l)CQ_z\sum_m a(m)E_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+l+m+1,k\right)$$
$$=CA_xE_x^{n}\left(i+\tfrac{1}{2},j,k\right)+CB_x\times$$
$$\left(\begin{array}{l}\dfrac{1}{\Delta y}\sum_l a(l)CP_zH_z^{n}\left(i+\tfrac{1}{2},j+l+\tfrac{1}{2},k\right)\\ -\dfrac{1}{\Delta z}\sum_l a(l)H_y^{n}\left(i+\tfrac{1}{2},j,k+l+\tfrac{1}{2}\right)\\ -\dfrac{1}{\Delta x\Delta y}\sum_l a(l)CQ_z\sum_m a(m)E_y^{n}\left(i+m+1,j+l+\tfrac{1}{2},k\right)\end{array}\right). \tag{6}$$

Observing Eq. (1), we can see that Eq. (1) cannot be solved directly because both sides of Eq. (1) contain unknown components. It needs the unknown $H_z^{n+1/2}$ component to compute the $E_x^{n+1/2}$ component, so we can substitute Eq. (2) into Eq. (1), and the equation for $E_x^{n+1/2}$ can be obtained as Eq. (6) by proper rearrangement.

In free space, the above formula can be simplified as:

$$E_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j,k\right)-\frac{\Delta t^2}{4\varepsilon\mu\Delta y^2}\times$$
$$\sum_l a(l)\sum_m a(m)E_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+l+m+1,k\right)$$
$$=E_x^{n}\left(i+\tfrac{1}{2},j,k\right)+$$
$$\frac{\Delta t}{2\varepsilon}\left(\begin{array}{l}\dfrac{1}{\Delta y}\sum_l a(l)H_z^{n}\left(i+\tfrac{1}{2},j+l+\tfrac{1}{2},k\right)\\ -\dfrac{1}{\Delta z}\sum_l a(l)H_y^{n}\left(i+\tfrac{1}{2},j,k+l+\tfrac{1}{2}\right)\\ -\dfrac{\Delta t}{2\mu\Delta x\Delta y}\sum_l a(l)\sum_m a(m)E_y^{n}\left(i+m+1,j+l+\tfrac{1}{2},k\right)\end{array}\right). \tag{7}$$

Similarly, we can obtain the equation for electric fields at the (n+1)th step:

$$E_x^{n+1}\left(i+\tfrac{1}{2},j,k\right)-$$
$$\frac{\Delta t^2}{4\varepsilon\mu\Delta z^2}\sum_l a(l)\sum_m a(m)E_x^{n+1}\left(i+\tfrac{1}{2},j,k+l+m+1\right)$$
$$=E_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j,k\right)+$$
$$\frac{\Delta t}{2\varepsilon}\left(\begin{array}{l}\dfrac{1}{\Delta y}\sum_l a(l)H_z^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+l+\tfrac{1}{2},k\right)\\ -\dfrac{1}{\Delta z}\sum_l a(l)H_y^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j,k+l+\tfrac{1}{2}\right)\\ -\dfrac{\Delta t}{2\mu\Delta x\Delta z}\sum_l a(l)\sum_m a(m)E_z^{n+\frac{1}{2}}\left(i+m+1,j,k+l+\tfrac{1}{2}\right)\end{array}\right). \tag{8}$$

Taking Eq. (6) as example, the left side of Eq. (6) can be expressed as:

$$\begin{array}{l}A_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j-5,k\right)+B_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j-4,k\right)+\\ C_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j-3,k\right)+D_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j-2,k\right)+\\ E_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j-1,k\right)+F_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j,k\right)+\\ G_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+1,k\right)+H_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+2,k\right)+\\ I_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+3,k\right)+J_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+4,k\right)+\\ K_j^eE_x^{n+\frac{1}{2}}\left(i+\tfrac{1}{2},j+5,k\right)=L_j^e\end{array}, \tag{9}$$

where

$$\left.\begin{aligned}
&A_j^e=-\frac{a^2(-3)\Delta t^2}{4\varepsilon\mu\Delta y^2},\ B_j^e=-\frac{2a(-3)a(-2)\Delta t^2}{4\varepsilon\mu\Delta y^2}\\
&C_j^e=-\frac{\left[2a(-3)a(-1)+a^2(-2)\right]\Delta t^2}{4\varepsilon\mu\Delta y^2}\\
&D_j^e=-\frac{\left[2a(-3)a(0)+2a(-2)a(-1)\right]\Delta t^2}{4\varepsilon\mu\Delta y^2}\\
&E_j^e=-\frac{\left[2a(-3)a(1)+2a(-2)a(0)+a^2(-1)\right]\Delta t^2}{4\varepsilon\mu\Delta y^2}\\
&F_j^e=1-\frac{\left[2a(-3)a(2)+2a(-2)a(1)+2a(-1)a(0)\right]\Delta t^2}{4\varepsilon\mu\Delta y^2}\\
&J_j^e=-\frac{\left[2a(-2)a(2)+2a(-1)a(1)+a^2(0)\right]\Delta t^2}{4\varepsilon\mu\Delta y^2}\\
&H_j^e=-\frac{\left[2a(-1)a(2)+2a(0)a(1)\right]\Delta t^2}{4\varepsilon\mu\Delta y^2}\\
&I_j^e=-\frac{\left[2a(0)a(2)+a^2(1)\right]\Delta t^2}{4\varepsilon\mu\Delta y^2}\\
&J_j^e=-\frac{2a(1)a(2)\Delta t^2}{4\varepsilon\mu\Delta y^2},\ K_j^e=-\frac{a^2(2)\Delta t^2}{4\varepsilon\mu\Delta y^2}
\end{aligned}\right\}. \tag{10}$$

The Eq. (9) can be written in a matrix form as follows:

$$\mathbf{AX}=\mathbf{Y}, \tag{11}$$

where the coefficients matrix is as follows:

$$\mathbf{A}=\begin{bmatrix}
\mathbf{a}_1 & \mathbf{b}_1 & 0 & \cdots & 0 & \mathbf{t}\\
\mathbf{c}_2 & \mathbf{a}_2 & \mathbf{b}_2 & \cdots & 0 & 0\\
0 & \mathbf{c}_3 & \mathbf{a}_3 & \mathbf{b}_3 & \cdots & 0\\
\vdots & 0 & \ddots & \ddots & \ddots & \vdots\\
0 & \vdots & \cdots & \mathbf{c}_{n-1} & \mathbf{a}_{n-1} & \mathbf{b}_{n-1}\\
\mathbf{s} & 0 & 0 & \cdots & \mathbf{c}_n & \mathbf{a}_n
\end{bmatrix}. \tag{12}$$

Matrix $\mathbf{A}$ is a block periodic tri-diagonal matrix, and $\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$, $\mathbf{t}$, $\mathbf{s}$ are all m order square matrices. And the form of coefficient matrix $\mathbf{A}$ in Eq. (11) can be written as Eq. (12). In this work, the PS-ADI-MRTD equations are obtained based on the D2 wavelet [14]-[16], and $m=5$. The elements of matrices $\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$, $\mathbf{t}$ and $\mathbf{s}$ are as follows:

$$\mathbf{a}_1=\begin{bmatrix}
F_1^e & G_1^e & H_1^e & I_1^e & J_1^e\\
E_2^e & F_2^e & G_2^e & H_2^e & I_2^e\\
D_3^e & E_3^e & F_3^e & G_3^e & H_3^e\\
C_4^e & D_4^e & E_4^e & F_4^e & G_4^e\\
B_5^e & C_5^e & D_5^e & E_5^e & F_5^e
\end{bmatrix}, \tag{13}$$

$$\mathbf{b}_1=\begin{bmatrix}
K_1^e & 0 & 0 & 0 & 0\\
J_2^e & K_2^e & 0 & 0 & 0\\
I_3^e & J_3^e & K_3^e & 0 & 0\\
H_4^e & I_4^e & J_4^e & K_4^e & 0\\
G_5^e & H_5^e & I_5^e & J_5^e & K_5^e
\end{bmatrix}, \tag{14}$$

$$\mathbf{c}_2=\begin{bmatrix}
A_6^e & B_6^e & C_6^e & D_6^e & E_6^e\\
0 & A_7^e & B_7^e & C_7^e & D_7^e\\
0 & 0 & A_8^e & B_8^e & C_8^e\\
0 & 0 & 0 & A_9^e & B_9^e\\
0 & 0 & 0 & 0 & A_{10}^e
\end{bmatrix}, \tag{15}$$

$$\mathbf{t}=\begin{bmatrix}
A_1^e & B_1^e & C_1^e & D_1^e & E_1^e\\
0 & A_2^e & B_2^e & C_2^e & D_2^e\\
0 & 0 & A_3^e & B_3^e & C_3^e\\
0 & 0 & 0 & A_4^e & B_4^e\\
0 & 0 & 0 & 0 & A_5^e
\end{bmatrix}\exp\left(-jk_yT_y\right), \tag{16}$$

$$\mathbf{s}=\begin{bmatrix}
K_{n-4}^e & 0 & 0 & 0 & 0\\
J_{n-3}^e & K_{n-3}^e & 0 & 0 & 0\\
I_{n-2}^e & J_{n-2}^e & K_{n-2}^e & 0 & 0\\
H_{n-1}^e & I_{n-1}^e & J_{n-1}^e & K_{n-1}^e & 0\\
G_n^e & H_n^e & I_n^e & J_n^e & K_n^e
\end{bmatrix}\exp\left(jk_yT_y\right). \tag{17}$$

It is defined that $\mathbf{p}_1$, $\mathbf{q}_1$ are m order square matrices, $\mathbf{q}_n=\mathbf{p}_1^{-1}\mathbf{t}$, $\mathbf{p}_n=\mathbf{s}\mathbf{q}_1^{-1}$, and $\mathbf{p}=\begin{bmatrix}\mathbf{p}_1 & 0 & \cdots & 0 & \mathbf{p}_n\end{bmatrix}^{\mathrm{T}}$, $\mathbf{q}=\begin{bmatrix}\mathbf{q}_1 & 0 & \cdots & 0 & \mathbf{q}_n\end{bmatrix}^{\mathrm{T}}$, matrix $\mathbf{A}$ can be written as $\mathbf{A}=\mathbf{B}+\mathbf{p}\mathbf{q}^{\mathrm{T}}$, where:

$$\mathbf{B}=\begin{bmatrix}
\mathbf{a}_1-\mathbf{p}_1\mathbf{q}_1 & \mathbf{b}_1 & 0 & \cdots & 0 & 0\\
\mathbf{c}_2 & \mathbf{a}_2 & \mathbf{b}_2 & \cdots & 0 & 0\\
0 & \mathbf{c}_3 & \mathbf{a}_3 & \mathbf{b}_3 & \cdots & 0\\
\vdots & 0 & \ddots & \ddots & \ddots & \vdots\\
0 & \vdots & \cdots & \mathbf{c}_{n-1} & \mathbf{a}_{n-1} & \mathbf{b}_{n-1}\\
\mathbf{s} & 0 & 0 & \cdots & \mathbf{c}_n & \mathbf{a}_n-\mathbf{p}_n\mathbf{q}_n
\end{bmatrix}. \tag{18}$$

Following the Sherman-Morrison-Woodbury formula, we can get [17]:

$$\mathbf{A}^{-1}=\left(\mathbf{B}+\mathbf{p}\mathbf{q}^{\mathrm{T}}\right)^{-1}=\mathbf{B}^{-1}-\mathbf{B}^{-1}\mathbf{p}\left(\mathbf{I}_m+\mathbf{q}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{p}\right)^{-1}\mathbf{q}^{\mathrm{T}}\mathbf{B}^{-1}. \tag{19}$$

Suppose that matrix $\mathbf{B}$ is invertible and then the necessary and sufficient condition for matrix $\mathbf{A}$ to be invertible is that the $\mathbf{I}_m+\mathbf{q}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{p}$ is invertible. Finally, $\mathbf{A}^{-1}$ can be solved as follows:

$$\mathbf{A}^{-1}=\left(\mathbf{c}_{ij}\right)_{n\times n},\ \mathbf{c}_{ij}=\begin{cases}\mathbf{g}_i\mathbf{h}_j-\boldsymbol{\psi}_i\boldsymbol{\zeta}_j, & i\le j\\ \mathbf{x}_i\mathbf{y}_j-\boldsymbol{\psi}_i\boldsymbol{\zeta}_j, & i>j\end{cases}. \tag{20}$$

And $\mathbf{c}_{ij}$ is calculated as follows. For the given invertible matrices $\mathbf{p}_1$ and $\mathbf{q}_1$, where $\mathbf{q}_n = \mathbf{p}_1^{-1}\mathbf{t}$, $\mathbf{p}_n = \mathbf{s}\mathbf{q}_1^{-1}$, one can have:

$$\mathbf{u}_1 = \mathbf{a}_1 - \mathbf{p}_1\mathbf{q}_1$$
$$\mathbf{l}_{i-1} = \mathbf{c}_i\mathbf{u}_{i-1}^{-1},\ \mathbf{u}_i = \mathbf{a}_i - \mathbf{l}_{i-1}\mathbf{b}_{i-1}\ (i = 2,3,\cdots,n-1)$$
$$\mathbf{u}_n = \mathbf{a}_n - \mathbf{p}_n\mathbf{q}_n - \mathbf{l}_{n-1}\mathbf{b}_{n-1}$$
$$\boldsymbol{\alpha}_n = \mathbf{a}_n - \mathbf{p}_n\mathbf{q}_n,\ \boldsymbol{\beta}_i = \boldsymbol{\alpha}_{i+1}^{-1}\mathbf{c}_{i+1},$$
$$\boldsymbol{\alpha}_i = \mathbf{a}_i - \mathbf{b}_i\boldsymbol{\beta}_i\ (i = n-1, n-2,\cdots,1).$$

Defining that $\mathbf{g}_1 = \mathbf{I}_m$, $\mathbf{x}_1 = \mathbf{I}_m$, one can have:

$$\mathbf{h}_1 = \boldsymbol{\alpha}_i^{-1},\ \mathbf{h}_i = -\mathbf{b}_{i-1}\mathbf{h}_{i-1}\boldsymbol{\alpha}_i^{-1}\ (i = 2,3,\cdots,n)$$
$$\mathbf{g}_n = (\mathbf{u}_n\mathbf{h}_n)^{-1},\ \mathbf{g}_i = -\mathbf{b}_i\mathbf{g}_{i+1}\mathbf{u}_i^{-1}\ (i = n-1, n-2,\cdots,2)$$
$$\mathbf{y}_1 = \boldsymbol{\alpha}_1^{-1},\ \mathbf{y}_i = -\boldsymbol{\beta}_{i-1}\mathbf{y}_{i-1}\ (i = 2,3,\cdots,n)$$
$$\mathbf{x}_n = (\mathbf{u}_n\mathbf{h}_n)^{-1},\ \mathbf{x}_i = -\mathbf{l}_i\mathbf{x}_{i+1}\ (i = n-1, n-2,\cdots,2)$$
$$\boldsymbol{\xi} = \mathbf{I}_m + (\mathbf{q}_1\mathbf{g}_1\mathbf{h}_1 + \mathbf{q}_n\mathbf{y}_n\mathbf{x}_1)\mathbf{p}_1 + (\mathbf{q}_1\mathbf{g}_1\mathbf{h}_n + \mathbf{q}_n\mathbf{y}_n\mathbf{x}_n)\mathbf{p}_n$$
$$\boldsymbol{\psi} = \boldsymbol{\xi}^{-1}(\mathbf{p}_1\mathbf{x}_1\mathbf{y} + \mathbf{p}_n\mathbf{y}_n\mathbf{g}),\ \boldsymbol{\zeta} = \mathbf{q}_1\mathbf{g}_1\mathbf{h}^{\mathrm{T}} + \mathbf{q}_n\mathbf{y}_n\mathbf{x}^{\mathrm{T}}$$
$$\mathbf{c}_{ij} = \begin{cases} \mathbf{g}_i\mathbf{h}_j - \boldsymbol{\psi}_i\boldsymbol{\zeta}_j, & i \le j \\ \mathbf{x}_i\mathbf{y}_j - \boldsymbol{\psi}_i\boldsymbol{\zeta}_j, & i > j \end{cases}.$$

The TF/SF boundary and the absorbing boundary condition (ABC) of the PS-ADI-MRTD algorithm are similar to that of the traditional MRTD algorithm.

III. HANDLIING OF THE PERIODIC BOUNDARY CONDITION

For a better understanding of this paper, Fig. 2 shows a top view of a unit in a periodic structure, where T_x and T_y are the cycle lengths along the *x*- and *y*-direction, respectively, Nx_{F}, Nx_{L} and Ny_{F}, Ny_{L} are the first and last grid points along the *x*- and *y*-direction in the computational domain, respectively.

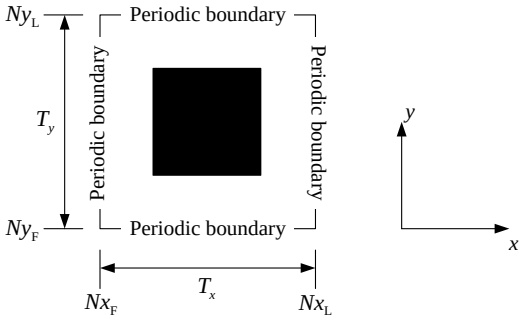


Fig. 2. Schematic diagram of the periodic structure unit.

Taking the electric fields as example, the periodic boundary condition of the MRTD can be expressed as follows:

$$E_x^n\left(i+\tfrac{1}{2}, Ny_{\mathrm{F}} \pm l, k\right) = E_x^n\left(i+\tfrac{1}{2}, Ny_{\mathrm{L}} \pm l, k\right)\exp\left(\mathrm{j}k_y T_y\right), \quad (21)$$
$$E_y^n\left(Nx_{\mathrm{F}} \pm l, j+\tfrac{1}{2}, k\right) = E_y^n\left(Nx_{\mathrm{L}} \pm l, j+\tfrac{1}{2}, k\right)\exp\left(\mathrm{j}k_x T_x\right), \quad (22)$$
$$E_z^n\left(i, Ny_{\mathrm{F}} \pm l, k+\tfrac{1}{2}\right) = E_z^n\left(i, Ny_{\mathrm{L}} \pm l, k+\tfrac{1}{2}\right)\exp\left(\mathrm{j}k_y T_y\right), \quad (23)$$
$$E_z^n\left(Nx_{\mathrm{F}} \pm l, j, k+\tfrac{1}{2}\right) = E_z^n\left(Nx_{\mathrm{L}} \pm l, j, k+\tfrac{1}{2}\right)\exp\left(\mathrm{j}k_x T_x\right). \quad (24)$$

The other set of equations for **H** can be obtained by duality. Where $l = 0,\ 1,\ 2\cdots$, it should be noted that the *l* is an effective support size for the MRTD basis function.

IV. NUMERICAL RESULTS

A. Scattering analysis of thin metal square array

In this section, a thin square metal plate array is calculated. As shown in Fig. 1, the edge length of a square metal plate is 0.5 cm, and the periodic unit size is $1\mathrm{cm}\times 1\mathrm{cm}$. The space is discretized by a mesh $\Delta x = \Delta y = \Delta z = \Delta = 0.05$ cm, the time step is $\Delta t_{\mathrm{CFL}} = \Delta / 3c$, and $\mathrm{CFLN} = \Delta t / \Delta t_{\mathrm{CFL}}$, where Δt_{CFL} is the time step limit defined by the stability condition of the traditional MRTD. $\mathrm{CFLN} = 3$ is used for the PS-ADI-MRTD method. The computational domain along the *z*-axis is truncated with absorbing boundary. The excitation is a TE wave with constant transverse wave-number (CTW) and the wave vector is on the *x*-*z* plane with $k_y = 0$. Therefore, the time domain expressions of the incident electric and magnetic fields can be expressed as follows:

$$E_y^{CTW} = \exp(-jk_x x) \times F^{-1}\left[\exp(jk_z(z-z_0))\exp(-\frac{k_0^2}{\sigma^2})\exp(-jt_0k_0c)\right], \quad (25)$$

$$H_x^{CTW} = \frac{1}{\eta_0}\exp(-jk_x x) \times F^{-1}\left[\exp(jk_z(z-z_0))\frac{k_z}{k_0}\exp(-\frac{k_0^2}{\sigma^2})\exp(-jt_0k_0c)\right], \quad (26)$$

where k_x denotes transverse wave-numbers, which are assumed to be constant numbers (independent of frequency). k_z is the normal wave-number, $k_0 = 2\pi f / c$. η_0 is the impedance of free space. The term $\exp(-k_0^2/\sigma^2)$ corresponds to a Gaussian pulse used to limit the bandwidth of the incident wave, and $\sigma = 950$. F^{-1} denotes the inverse Fourier transform. The frequency range here is from 1.0 GHz to 40GHz, then the k_x range is from 0.0 to 838rad/m.

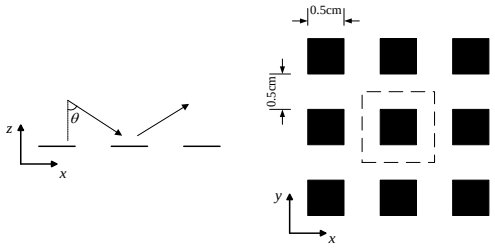


Fig. 1. Thin square metal sheet array.

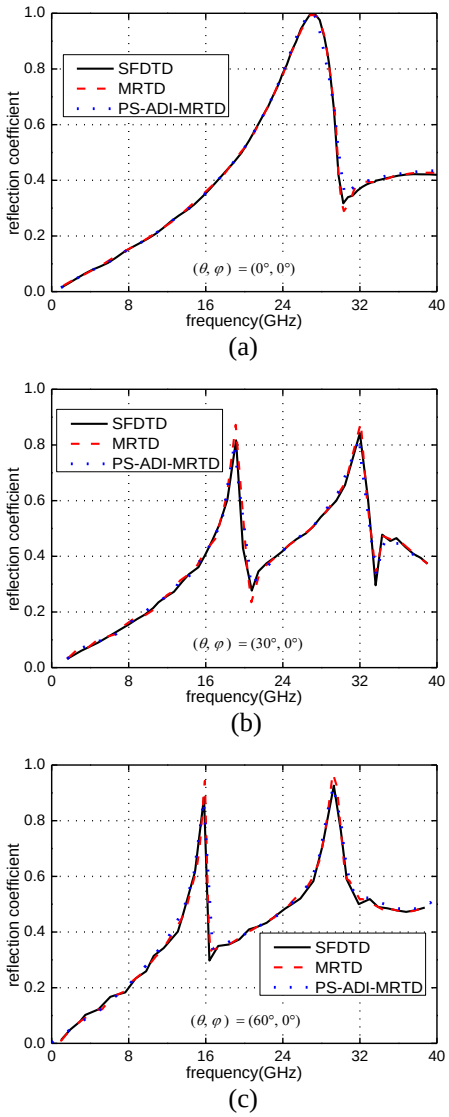


Fig. 2. The reflection coefficients with different incident angles.

Figure 2 shows the calculated reflection coefficients versus frequencies at different incident angles. As can be seen from Fig. 2, the numerical results of the proposed method are in good agreement with that of the MRTD and SFDTD, which verifies the accuracy of the PS-ADI-MRTD method. In addition, as shown in Table 1, numerical results also validate that although the PS-ADI-MRTD method uses more memory, it can save computing time by more than 26% compared with the traditional SFDTD and MRTD.

Table 1: CPU time and memory for different schemes

Schemes	CPU Time/s	CPU Memory/MB
SFDTD	3563.25	8.10
MRTD	1347.56	2.94
PS-ADI-MRTD	973.53	4.31

B. Application in frequency selective surface analysis

Frequency selective surface (FSS) is a kind of periodic structure which is widely used. It is generally composed of a certain number of passive resonant elements arranged in a specific way. Its main feature is that it can filter electromagnetic waves at different frequencies, incidence angles and polarization states. Two kinds of common FSS structures are periodically arranged sheet metal (patch type) and periodically opened sheet metal (aperture type). Generally speaking, in a certain frequency band near the resonance frequency, the former exhibits total reflection, while the latter exhibits total transmission [18]. In this section, the scattering characteristics in frequency domain are analyzed for the two types of FSS structures.

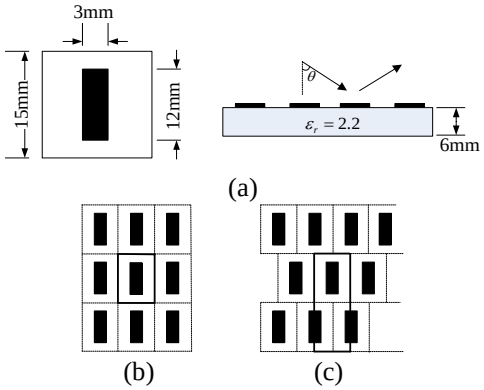


Fig. 3. Patch FSS structure and two array modes.

Figure 3 (a) is a periodic unit of a patch-type FSS structure and its placement on the substrate medium (black part is metal). The relative conductivity of the medium is $\varepsilon_r = 2.2$ and the thickness is 6 mm. Figure 3 (b) and Fig. 3 (c) are two kinds of array modes of FSS. The thickened wireframes in the figure are periodic units in two cases respectively. The setting of the constant transverse incident wave is also the same as that in Section III.A.

Figure 4 is the calculation result of two arrays under vertical incidence. It can be seen that the translation of the structure results in a slight shift of the reflection characteristic curve to the high frequency. The resonance frequencies before and after the translation are 9.5 GHz and 9.7 GHz respectively, and the second resonance frequency point appears in the array arrangement mode (b) at about 16.5 GHz.

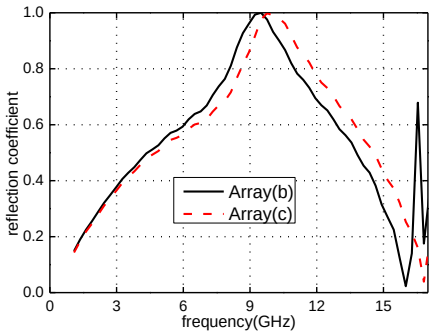


Fig. 4. Frequency characteristics of two array modes in vertical incidence.

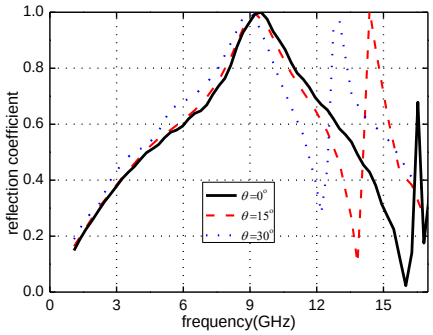


Fig. 5. Frequency characteristics of array mode (b) at different incidence angles.

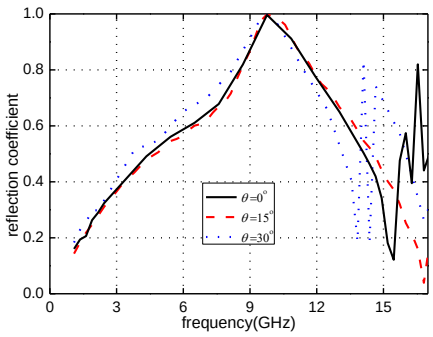


Fig. 6. Frequency characteristics of array mode (c) at different incidence angles.

Figure 5 and Fig. 6 show the results of two array modes when the incidence angle is 0°, 15° and 30°, respectively. It can be seen that the variation law of the two graphs is the same, and the reflection characteristic curve moves to the low frequency end with the increase of the incident angle. The first resonance frequency point appearing under the three incident angles has little difference, and the characteristic curve before the frequency point is also similar. As shown in Fig. 5, the second resonance frequencies of the array arrangement (b) at 15 and 30 degrees are 14.4 GHz and 13 GHz, respectively. From Fig. 6, it can be seen that the second and third resonance frequencies appear in the calculated frequency band after the array translation, such as the two consecutive resonance frequencies of 14 GHz and 14.6 GHz at the incident angle of 30 degrees, respectively.

Using the same method, we also calculated the frequency domain scattering characteristics of the aperture-type FSS structure shown in Fig. 7.

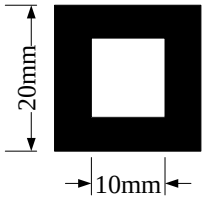
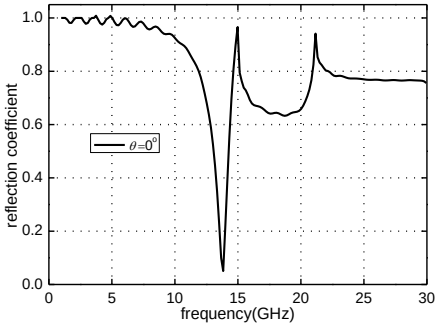
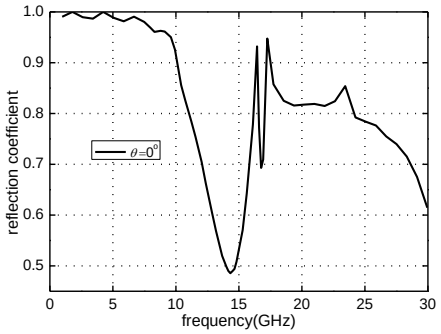


Fig. 7. Periodic cell of aperture-type FSS.

The calculated results at incident angles 0 degrees and 30 degrees are shown in Fig. 8 (a) and Fig. 8 (b) respectively. It can be seen that the resonant frequency point appears at about 13.8 GHz at vertical incidence, when the reflectivity is near zero and the corresponding transmittance is 1, which is consistent with the full transmission characteristics of the aperture FSS discussed above. When the incident angle increases, the resonant frequency is about 14.3GHz, and the reflectivity is about 0.48. It can also be seen from the figure that with the further increase of frequency, the frequency reflection characteristics of the structure are more complex.



(a)



(b)

Fig. 8. Frequency characteristics of the aperture-type FSS.

V. CONCLUSION

In this paper, the spectral technique is applied to ADI-MRTD method for periodic structure calculation, and the PS-ADI-MRTD method is obtained. The inverse process of block periodic tridiagonal matrix is introduced and the numerical results verify the effectiveness and efficiency of the proposed PS-ADI-MRTD method. In the calculation case, the CPU time saveing is about 26%. Moreover, we apply this algorithm to analyze the frequency domain scattering characteristics of FSS and obtain the expected results, which further confirms the effectiveness of PS-ADI-MRTD algorithm.

REFERENCES

[1] M. Krumpholz and L. P. B. Katehi, "New prospects for time domain analysis," *IEEE Microwave Guid. Wave Lett.*, vol. 5, no.11, pp. 382-384, Dec. 1995.

[2] M. Krumpholz and L. P. B. Katehi, "MRTD: New time-domain schemes based on multiresolution analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 44, no. 4, pp. 555-561, Apr. 1996.

[3] Y. W. Liu, Y. W. Chen, P. Zhang, and X. Xu, "Implementation and application of the spherical MRTD algorithm," *Progress In Electromagnetics Research*, vol. 139, pp. 577-597, 2013.

[4] Y. W. Liu, Y. W. Chen, and P. Zhang, "Parallel implementation and application of the MRTD with an efficient CFS–PML," *Progress in Electromagnetics Research*, vol. 143, pp. 223-242, 2013.

[5] P. Zhang, Y. W. Liu, S. Qiu, and B. Yang "CPML and quasi-CPML for cylindrical MRTD method," *Progress In Electromagnetics Research B*, vol. 61, pp. 17-30, 2014.

[6] C. Zhizhang and J. Zhang, "An unconditionally stable 3-D ADI-MRTD method free of the CFL stability condition," *IEEE Microwave and Wireless Components Letters*, vol. 11, no. 8, pp. 349-351, 2001.

[7] X. Dardenne and C. Craeye, "Method of moments simulation of infinitely periodic structures combining metal with connected dielectric objects," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 8, pp. 2372-2380, 2008.

[8] L. E. R. Petersson and J.-M. Jin, "A two-dimensional time-domain finite element formulation for periodic structures," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 1, pp. 12-19, 2005.

[9] S. Wang, J. Chen, and P. Ruchhoeft, "An ADI-FDTD method for periodic structures," *IEEE Transactions on Antennas and Propagation*, vol. 53, pp. 2343-2346, 2005.

[10] G. Singh, E. L. Tan, and Z. N. Chen, "Efficient complex envelope ADI-FDTD method for the analysis of anisotropic photonic crystals," *IEEE Photonics Technology Letters*, vol. 23, pp. 801-803, 2011.

[11] A. Amir and R. S. Yahya, "Spectral FDTD: A novel technique for the analysis of oblique incident plane wave on periodic structures," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 6, pp. 1818-1825, 2006.

[12] F. Yang, A. Elsherbeni, and J. Chen, "A hybrid spectral-FDTD/ARMA method for periodic structure analysis," *IEEE Antennas and Propagation Society International Symposium*, pp. 3720-3723, 2007.

[13] Y. F. Mao, B. Chen, H. Q. Liu, J. L. Xia ,and J. Z. Tang, "A hybrid implicit-explicit spectral FDTD scheme for oblique incidence problems on periodic structures," *Progress In Electromagnetics Research*, vol. 128, pp. 153-170, 2012.

[14] Y. W. Cheong, Y. M. Lee, K. H. Ra, J. G. Kang, and C. C. Shin, "Wavelet-Galerkin scheme of time-dependent inhomogeneous electromagnetic problems," *IEEE Microwave Guid, Wave Lett.*, vol. 9, no. 8, pp. 297-299, Aug. 1999.

[15] M. Fujii and W. J. R. Hoefer, "Dispersion of time domain wavelet Galerkin method based on Daubechies' compactly supported scaling functions with three and four vanishing moments," *IEEE Microwave Guid. Wave Lett.*, vol. 10, no. 4 pp. 125-127, Apr. 2000.

[16] I. Daubechies, *Ten Lectures on Wavelets*. SIAM, Philadelphia, PA, 1992.

[17] Y. Du, Q. Lu, and Z. Xu, "New algorithm for inversing block periodic tridiagonal matrices," *Computer Engineering and Applications*, vol. 48, no. 17, pp. 41-43, 2012.

[18] Y. Rahmat-Samii and H. Mosallaei, "Electromagnetic band-gap structures: Classsification, characterization and applications," *IEE-ICAP Symposium, Manchester, United Kingdom*, pp. 560-564, Apr. 17-20, 2001.

Robust Optimization of Electromagnetic Design Using Stochastic Collocation Method

Gang Zhang [1], Ruihuan Zhu [1], Jinjun Bai [2], and Xiyuan Peng [1]

[1] School of Electrical Engineering and Automation
Harbin Institute of Technology, Harbin, 150006, China
zhang_hit@hit.edu.cn,18S106121@stu.hit.edu.cn, pxy@hit.edu.cn

[2] College of Marine Electrical Engineering
Dalian Maritime University, Dalian, 116026, China
baijinjun@dlmu.edu.cn

Abstract ─ The way to handle the uncertainty of design parameters has attracted wide attention in the optimization of electromagnetic designs. The Monte Carlo Method works well when dealing with uncertainties but it consumes too much time and computational resources. This paper proposes a computationally efficient way to achieve robust optimization based on Stochastic Collocation Method and the TEAM 22 problem is used as a verification example. It is demonstrated that the approach combining Stochastic Collocation Method and a genetic algorithm provides high computational efficiency, without losing accuracy compared with the Monte Carlo Method.

Index Terms ─ Monte Carlo Method, robust, Stochastic Collocation Method, TEAM 22, uncertainty.

I. INTRODUCTION

It is a common knowledge that the performance of the electromagnetic (EM) design is influenced by uncertainties [1]. The uncertainties may be caused by, for example, manufacturing tolerances in geometric variables or the inevitable variation of material parameters. These uncertainties will jeopardize the robustness of the design. Robustness refers to the insensitivity of the final optimization results to parameter perturbations [4]. A robust design should not be very sensitive to slight changes in its design parameters as this would either seriously impact production costs or make the physical machine behave differently from its optimized computer simulation model [5].

In response to this requirement, many scholars have used a variety of methods to obtain a robust device or control method. Some examples of the state-of-the-art optimization methods used in the field comprise a new algorithm based on the Climb method [4], a possibility-based optimal design algorithm [5], a surrogate modeling technique based on a second-order equation [6], and the space-time kriging surrogate model [8].

The previous methods require a lot of time and computational resources to obtain a robust solution, which significantly affect computational efficiency of EM optimization [9] .Therefore, some measures are taken such as evaluating the robustness of both performance and constraints under uncertainty by the worst-case optimization [9], reusing the global surrogate model to perform sensitivity analyses of generated designs [10], and proposing an efficient reliability-based robust design optimization method [11]. However, these methods more or less show limitations when they are applied to other optimization problems.

To improve the numerical efficiency of EM optimization while maintaining design robustness, an optimization strategy based on Stochastic Collocation Method (SCM) [12] is proposed in this paper. The basic concepts of SCM are given in Section II. Section III shows the process of obtaining the robust optimal solution based on SCM. Section IV introduces the TEAM 22 problem [17] and compares the robustness and time consumption of the optimization techniques, with the results obtained by the different methods. Finally, the paper's conclusions are drawn in Section V.

II. OUTLINE OF THE STOCHASTIC COLLOCATION METHOD

It can be helpful to apply uncertainty analysis methods to computational electromagnetics (CEM) [20] in order to take account of practical complexity and unpredictability within a simulation. To this end, design parameters of EM simulation are presented by random variables with properly assigned distributions.

The Stochastic Collocation method (SCM) is a popular choice for the stochastic processing of complex systems where well-established deterministic codes exist. SCM is supposed to be universal because it only

Submitted On: November 26, 2019
Accepted On: April 12, 2020

deals with the input and output of the problem without changing the solver itself. The applicability of SCM method is not affected by the complexity of the original problem so long as reliable deterministic solver is developed. By utilizing the SCM method, the relationship between the uncertainty of output and the input variables is approximated by the sum of some specific polynomials. One way is to use a Lagrange interpolation approach that is given by (1):

$$y(x) = Lag(f(x)) = \sum_{j=0}^{n} f(x_j) l_j(x), \tag{1}$$

where $y(x)$ is a polynomial approximation of the true solution $f(x)$. n is the number of collocation points. x_j and $f(x_j)$ stand for the collocation points and the corresponding deterministic solutions of these points, respectively. $l_j(x)$ are the Lagrange interpolation polynomials structured by the collocation points:

$$l_j(x) = \prod_{i=0, i\neq j}^{n} \frac{(x - x_i)}{(x_j - x_i)}. \tag{2}$$

As for the SCM, the collocation points are given by the zero points of the generalized Polynomial Chaos [12]. The orthogonal polynomial basis is selected according to the probability distribution of the random variables, as shown in Table 1. Specially, if the variables are multidimensional, the interpolating points are the tensor product form of interpolating points in every dimension [16]. The accuracy of the SCM had been elaborated in [22]. It was shown that the longer each single CEM simulation lasted, the more efficient and thus more desirable SCM could be.

Table 1: The correspondence between the type of generalized Polynomial Chaos and Random Variables

Random Variables	Wiener-Askey Chaos	Support
Gaussian	Hermite-chaos	$(-\infty, +\infty)$
Gamma	Laguerre-chaos	$[0, +\infty)$
Beta	Jacobi-chaos	$[a, b]$
Uniform	Legendre-chaos	$[a, b]$

III. SCM BASED ROBUST OPTIMAL DESIGN

The traditional optimal design problem is constructed as:

$$\begin{aligned} &\min\ f(\boldsymbol{P}_d) \\ &s.t.\ g_i(\boldsymbol{P}_d) \le 0, \quad i = 1, ..., m, \end{aligned} \tag{3}$$

where $f(\boldsymbol{P}_d)$ is the objective function for design variable set $\boldsymbol{P}_d$; $g_i(\boldsymbol{P}_d)$ are the constraint functions for i=1,…,m.

When the uncertainties are taken into account, the random design variable set $\boldsymbol{P}$ is defined as:

$$\boldsymbol{P} = \overline{\boldsymbol{P}} + \boldsymbol{\xi},$$

where $\overline{\boldsymbol{P}}$ is the mean value of $\boldsymbol{P}$ according to the statistical definition, while $\boldsymbol{\xi}$ is the random variable whose distribution can be assumed to be uniform in order to obtain a robust solution to the optimization problem. Thus, $\boldsymbol{P}$ can be redefined by normalization as:

$$\boldsymbol{P} = \overline{\boldsymbol{P}} + \boldsymbol{\eta} \cdot x, \tag{4}$$

where $\boldsymbol{\eta}$ is half of the range of the distribution and x a random number between [-1,1].

The incorporation of the robustness analysis with formulation (4) incurs high computational time. For the traditionally used Monte Carlo method (MCM), a serious computational burden is imposed by the required large sample size, as well as the iterative nature of the design optimization process. In contrast, the SCM is similar to MCM in the sense that it involves only the solution of a sequence of deterministic calculations at given collocation points in the stochastic space [23].

By applying the SCM to robust optimal design problem, $f(\boldsymbol{P})$ can be approximated by Lagrange interpolation polynomial according to (1):

$$f(\boldsymbol{P}) \approx y(\boldsymbol{P}) = \sum_{i=1}^{n} f(\boldsymbol{P}_i) l_i(\boldsymbol{P}), \tag{5}$$

where $\boldsymbol{P}_i$ (i=1,2,…,n) is the collocation points for univariate (or tensor product form of interpolating points for multivariate) problems, with a similar form to (4):

$$\boldsymbol{P}_i = \overline{\boldsymbol{P}} + \boldsymbol{\eta} \cdot x_i.$$

Generally, the zero points of generalized Polynomial Chaos in Table I are chosen to be x_i. For instance, the Legendre polynomials are orthogonal with respect to the uniform distribution. Its expression is as follows:

$$\begin{aligned} &L_0(x) = 1, \\ &L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left\{ \left(x^2 - 1 \right)^n \right\}, \quad n = 1, 2, ... \end{aligned} \tag{6}$$

Let the zero points of the n-dimensional Legendre polynomial be $x_1, x_2, ..., x_n$, and the Lagrange basis polynomial $l_i(x)$ are given by (2).

After the $y(\boldsymbol{P})$ is constructed (equation (5)), the mean of the obtained n values is defined as the objective function value of $\boldsymbol{P}$ as MCM does. It is denoted as:

$$\hat{f}_S(\boldsymbol{P}) = \frac{1}{n} \sum_{i=1}^{n} y(\boldsymbol{P}_i) \tag{7}$$

The SCM is not subject to the number of sampling points but to the number of variables. For problems where the solution is a smooth function of the random input variables and the dimension of the stochastic space is moderate, SCM has been shown to converge much faster than MCM [23].

IV. APPLICATION TO TEAM 22

The TEAM workshop problem 22 [14-16] is an optimization case of the Superconducting Magnetics

Energy Storage (SMES) that has been used as a benchmark problem in magneto statics. The TEAM 22 system is composed of two coils with opposite current densities. The configuration of the TEAM 22 is presented in Fig. 1, and its parameters are summarized in Table 2. The goal of TEAM 22 is to find the best configuration in SMES device to maintain the stored energy while minimizing the stray field. The stray field is represented by magnetic flux density $\boldsymbol{B}_{\text{stray}}$ and it is evaluated in 21 equidistant points marked on lines *a* and *b* in Fig. 1:

$$\boldsymbol{B}^2_{\text{stray}} = \frac{\sum_{i=1}^{21} \boldsymbol{B}^2_{\text{stray},i}}{21}. \tag{8}$$

Furthermore, to keep the superconductivity characteristic, the restriction is given by the inequality (9):

$$\|J\| \le (-6.4\|\boldsymbol{B}_{\max}\| + 54) \ \ (\text{A/mm}^2). \tag{9}$$

Since the current density of both coils is fixed in the TEAM 22, the inequality (9) can also be expressed as (10):

$$\|\boldsymbol{B}_{\max}\| \le 4.92\text{T}. \tag{10}$$

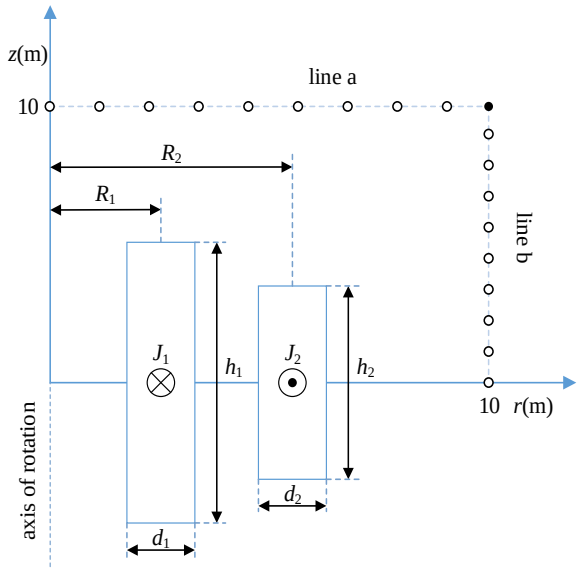


Fig. 1. The configuration of the TEAM 22.

When the robustness is taken into account in the optimization of TEAM 22, it is assumed that the adjustable parameters R_2, h_2, d_2 in the outside coil suffer undesirable and unavoidable presence of the uncertainties. The goals of the robustness-considered case remain the same as the classical problem. By defining the design variables set $\boldsymbol{P}=\{R_2, h_2, d_2\}$, the robust TEAM 22 problem can be formulated with the objectives and the restriction above as:

$$\min f(\boldsymbol{P}) = \omega_1 \frac{\boldsymbol{B}^2_{\text{stray}}}{\boldsymbol{B}^2_{\text{norm}}} + \omega_2 \frac{|E - E_0|}{E_0}, \tag{11}$$

$$s.t. \ \ \|J\| \le (-6.4\|\boldsymbol{B}_{\max}\| + 54)(\text{A/mm}^2)$$

where $\boldsymbol{B}_{\text{norm}}$=200μT; $\boldsymbol{B}_{\max}$ is the maximum magnetic flux density; E is the energy that is actually stored in the designed device; E_0 is the target stored energy with a fixed value 180MJ; ω_1, ω_2 are the barycentric weights. Take ω_1=0.001, ω_2=1 in this case to keep the relative value of the leakage flux in the same order of magnitude as the relative error of the stored energy:

$$\min f(\boldsymbol{P}) = 0.001 \times \frac{\boldsymbol{B}^2_{\text{stray}}}{\boldsymbol{B}^2_{\text{norm}}} + \frac{|E - E_0|}{E_0}.$$

$$s.t. \ \ \|B_{\max}\| \le 4.92\text{T}$$

The genetic algorithm (GA) is used to optimize the TEAM22, and the robustness is achieved by using MCM and SCM respectively at the fitness function which is the criterion for selection operators. The process of finding the optimal solution is shown in Fig. 2. Following the process of survival of the fittest in nature, the objective function is chosen as a selection criterion. Once the genetic representation and the fitness function are defined, the GA proceeds to initialize a population of solutions and then to improve it through repetitive application of the mutation, single point crossover, inversion and selection operators until a termination condition has been reached. The population is set to 200, and the termination condition is set to iterate 200 generations. The workstation used in this paper is Dell T7610 with Intel(R) Xeon(R) E5-2687W v2 3.4GHz and 128G RAM.

Table 2: Team 22 parameters

	R_1[m]	R_2[m]	h_1[m]	h_2[m]	d_1[m]	d_2[m]	J_1[A/mm^2]	J_2[A/mm^2]
Min	-	2.6	-	0.408	-	0.1	-	-
Max	-	3.4	-	2.2	-	0.4	-	-
Step	-	0.01	-	0.007	-	0.003	-	-
Fixed	2.0	-	1.6	-	0.27	-	22.5	-22

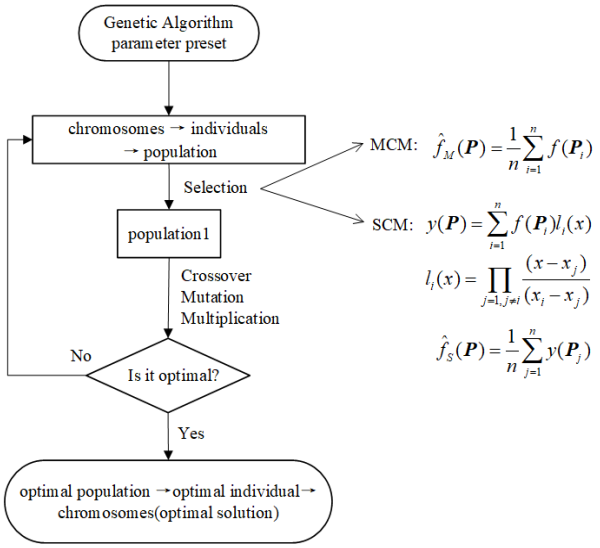


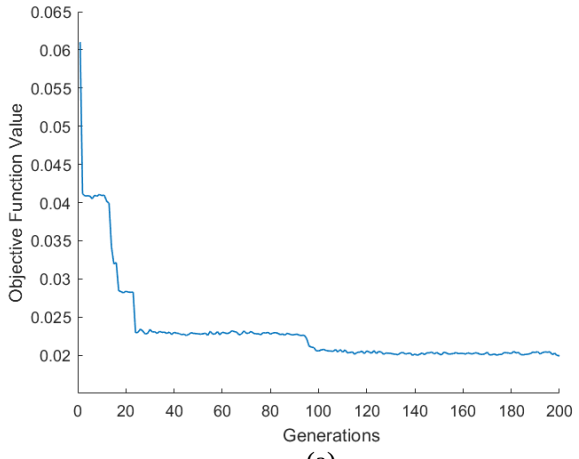
Fig. 2. The process of finding the optimal solution through GA.

A. Robustness achieved by MCM

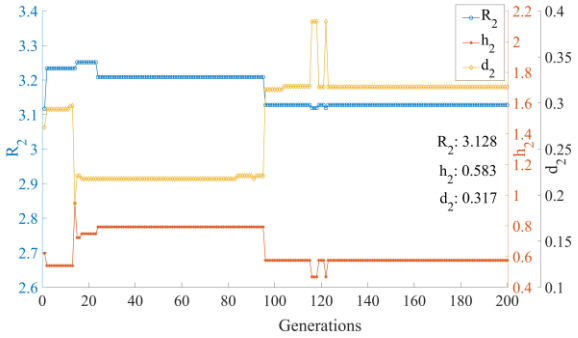
Let $\boldsymbol{\eta}=\{\eta_{R_2},\eta_{h_2},\eta_{d_2}\}=\{0.01,0.007,0.003\}$ be the set of variable step values given by Table 2. 100 points are randomly sampled in the range of $\overline{\boldsymbol{P}}\pm\boldsymbol{\eta}$ and denoted as $\boldsymbol{P}_1,\boldsymbol{P}_2,\ldots,\boldsymbol{P}_{100}$. Assume the notation $\hat{f}_M(\boldsymbol{P})$ for the mean of objective function value of the selected 100 points. Its expression is given by (12):

$$\hat{f}_M(\boldsymbol{P})=\frac{1}{100}\sum_{i=1}^{100}f(\boldsymbol{P}_i). \tag{12}$$

Every individual is treated with above process in the GA. And the result of optimization is shown in Fig. 3. The optimal solution is $\boldsymbol{P}$ = {3.128, 0.583, 0.317}.



(a)



(b)

Fig. 3. The result of optimization with robustness considered by MCM: (a) objective value and (b) parameters' value.

B. Robustness achieved by SCM

The cubic polynomials are selected in SCM to diminish the deviation of the interval. The three-dimensional form of the Lagrange interpolation formula is given by (13):

$$f(x,y,z)\approx y(x,y,z)=\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}f(x_i,y_j,z_k)l_k(z)l_j(y)l_i(x). \tag{13}$$

As for the TEAM 22, the interpolation formulation is formed as:

$$y(\boldsymbol{P})=\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}f(R_{2(i)},h_{2(j)},d_{2(k)})l_k(z)l_j(y)l_i(x), \tag{14}$$

where

$$\begin{cases}R_{2(i)}=\overline{R_2}+\eta_{R_2}\cdot x_i\\ h_{2(j)}=\overline{h_2}+\eta_{h_2}\cdot y_j\\ d_{2(k)}=\overline{d_2}+\eta_{d_2}\cdot z_k\end{cases}, \tag{15}$$

with x, y, z random numbers between [-1,1]. The interpolating points are $\{I_1,I_2,I_3\}=\left\{-\frac{\sqrt{15}}{5},0,\frac{\sqrt{15}}{5}\right\}$ where I can be x, y or z, and the tensor product form is $\left\{-\frac{\sqrt{15}}{5},0,\frac{\sqrt{15}}{5}\right\}\otimes\left\{-\frac{\sqrt{15}}{5},0,\frac{\sqrt{15}}{5}\right\}$. Thus, equation (14) can provide the answer.

Accordingly, 100 values of x, y and z are each randomly taken and substituted into (14) to obtain the corresponding objective function value .The notation $\hat{f}_S(\boldsymbol{P})$ is defined as the mean of the 100 obtained values

to represent the objective function value of $\boldsymbol{P}$ given by (16):

$$\hat{f}_S(\boldsymbol{P}) = \frac{1}{100}\sum_{i=1}^{100} y(\boldsymbol{P}_i). \tag{16}$$

The result of optimization is shown in Fig. 4. The optimal solution is $\boldsymbol{P}$ = {3.221,0.857,0.206}.

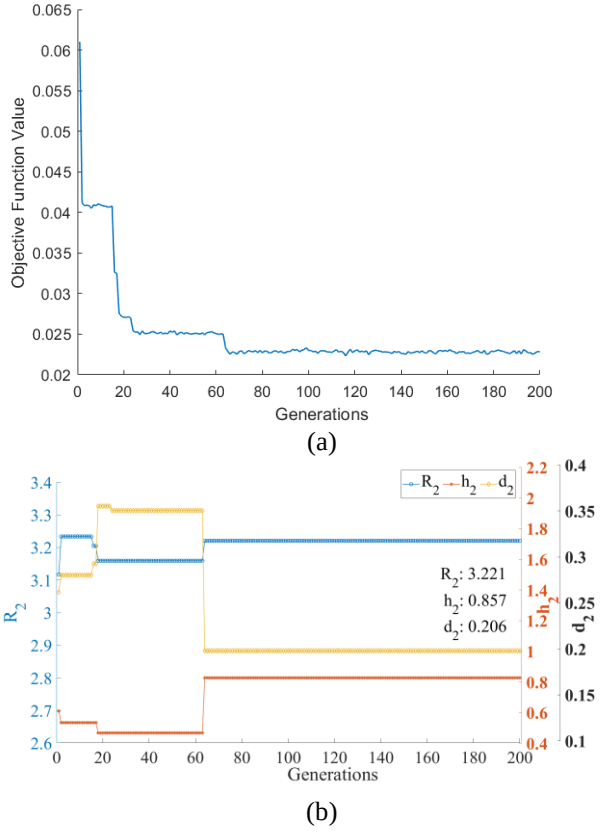


Fig. 4. The result of optimization with robustness considered by SCM: (a) objective value and (b) parameters' value.

C. Results and discussion

The rate of change in function value in the robust interval $\boldsymbol{D} = 0.6\boldsymbol{\eta}$ is defined as:

$$\delta = \frac{std(f)}{mean(f)},$$

where *std*(*f*) is the standard deviation of objective function *f* in the input parameters' interval while *mean*(*f*) is the mean of *f*.

The "normal" solution that does not consider robustness is also calculated. And it is compared with the solutions obtained above, as outlined in Table 3. Although the "normal" method obtains the smallest objective function value, the robustness of results is much lower than the other two methods.

The δ value that does not consider robustness is the largest when subjected to uncertainty perturbations, suggesting that the EM device considering robustness is more stable. In contrast, MCM and SCM reaches the same level of robustness. At the same time, it is noted that the time required by the optimal solution considering robustness is greatly increased. But the efficiency of achieving robustness by using SCM is improved by nearly 72.06%. Thus the SCM is an effective technique in terms of accuracy and computational efficiency.

For the SCM method, the order of the interpolating polynomials is set to be cubic in this paper. The cubic polynomials can lead to convergence considering both computational efficiency and accuracy according to [19]. To further justify this setting, several orders of interpolating polynomial are applied in the SCM based optimization. The results are compared with that of MCM, as shown in Fig. 5. The coincidence degree between probability distributions of MCM and SCM for different orders of interpolation is calculated, as outlined in Table 4, which means the accuracy of SCM is improved with the rise of order. Meanwhile, the consumed time Tc increases in the form of formula (17) with the rise of interpolating order where n stands for the number of order and p for the number of parameters (p is set to 3 in our case). If the time taken for two-point fitting is assumed for one unit, the relative times consumed by other orders are outlined in Table 4:

$$T_c = n^p. \tag{17}$$

It is noted that the accuracy of SCM decreases when interpolating order is higher than five. It is attributed to oscillation at the edges of an interval when using polynomial interpolation of high degree over a set of equispaced interpolation points, which is known as Runge's phenomenon. Therefore, the cubic order is selected to balance accuracy in the approximation of the objective function against Runge's phenomenon and computational requirements.

Table 3: Comparison of robust optimization results obtained by different methods

Method	*R*[m]	*h*[m]	*d*[m]	δ	Consumed Time[s]
Normal	3.127	0.548	0.336	7.58%	40078.79
MCM	3.128	0.583	0.317	6.85%	4709168.56
SCM	3.221	0.857	0.206	5.04%	1315861.82

Table 4: Comparison of robust optimization results

Orders of interpolation	2	3	4	5
Coincidence degree of probability distributions (taking the probability distribution of MCM as reference)	80.16%	90.41%	93.98%	95.46%
T_c (taking the time of two-point interpolation as reference)	1	3.375	8	15.625

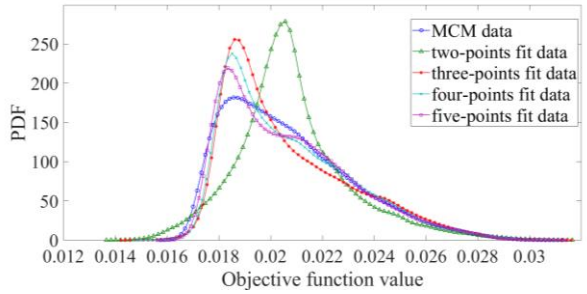


Fig. 5. Comparison of accuracy for different orders of interpolating polynomial.

It is indicated by (13) that the calculation process of SCM is related to the number of variables. It may take much more time to obtain a robust solution when the number of variables increases, which is known as the 'curse of dimensionality' problem. Therefore, SCM needs improvements and can be considered in combination with dimension reduced method [24] to solve the problem of multivariate uncertainty, which is an interesting issue to be investigated in future.

V. CONCLUSION

The uncertainties which exist in design variables and design process are taken into account in the optimization of EM devices. The SCM is applied to the robustness optimization by combing the GA algorithm, which exhibits comparable calculation accuracy to MCM with much less consumed time. For the TEAM 22 optimization, the computational efficiency is improved by nearly 72.06% using SCM. As a compromise between accuracy and efficiency, the order of interpolating polynomials for the SCM method is set to cubic. Meanwhile, the 'curse of dimensionality' problem of SCM caused by multivariate uncertainty still needs further investigation.

REFERENCES

[1] H. Li, I. Hussain, Y. Wang, and Q. Cao, "Uncertainty analysis of reflection coefficient for a coating with random flaws using adaptive mesh and DGTD method," *Applied Computational Electromagnetics Society Journal*, vol. 33, no. 5, 2018.

[2] J. Bai, G. Zhang, L. Wang, and T. Wang, "Uncertainty analysis in EMC simulation based on improved method of moments," *Applied Computational Electromagnetics Society Journal*, vol. 31, no. 1, 2016.

[3] L. Codecasa, L. Di Rienzo, K. Weise, and J. Haueisen, "Uncertainty quantification in transcranial magnetic stimulation with correlation between tissue conductivities," *2017 International Applied Computational Electromagnetics Society Symposium-Italy (ACES), IEEE*, 2017.

[4] J.-G. Lee, N.-W. Hwang, H.-R. Ryu, H.-K. Jung, and D.-K. Woo, "Robust optimization approach applied to permanent magnet synchronous motor," *IEEE Transactions on Magnetics*, vol. 53, no. 6, pp. 1-4, 2017.

[5] Z. Ren, S. He, D. Zhang, Y. Zhang, and C.-S. Koh, "A possibility-based robust optimal design algorithm in preliminary design stage of electromagnetic devices," *IEEE Transactions on Magnetics*, vol. 52, no. 3, pp. 1-4, 2016.

[6] G. Bramerdorfer, "Computationally efficient tolerance analysis of the cogging torque of brushless PMSMs," *IEEE Transactions on Industry Applications*, vol. 53, no. 4, pp. 3387-3393, 2017.

[7] G. Bramerdorfer, J. A. Tapia, J. Pyrhonen, and A. Cavagnino, "Modern electrical machine design optimization: Techniques, trends, and best practices," *IEEE Transactions on Industrial Electronics*, 2018.

[8] J. Jang, et al., "Space-time kriging surrogate model to consider uncertainty of time interval of torque curve for electric power steering motor," *IEEE Transactions on Magnetics*, vol. 54, no. 3, pp. 1-4, 2018.

[9] S. Xiao, Y. Li, M. Rotaru, and J. K. Sykulski, "Considerations of uncertainty in robust optimisation of electromagnetic devices," *International Journal of Applied Electromagnetics and Mechanics*, vol. 46, no. 2, pp. 427-436, 2014.

[10] G. Bramerdorfer and A.-C. Zăvoianu, "Surrogate-based multi-objective optimization of electrical machine designs facilitating tolerance analysis," *IEEE Transactions on Magnetics*, vol. 53, no. 8, pp. 1-11, 2017.

[11] B. Kang, K. Choi, and D-H. Kim, "An efficient serial-loop strategy for reliability-based robust optimization of electromagnetic design problems," *IEEE Transactions on Magnetics*, vol. 54, no. 3, pp. 1-4, 2018.

[12] J. Silly-Carette, D. Lautru, A. Gati, M. Wong, J. Wiart, and V. F. Hanna, "Determination of the uncertainty on the specific absorption rate using the stochastic collocation method and the FDTD," in *2008 IEEE Antennas and Propagation Society*

International Symposium, pp. 1-4, 2008.

[13] J. Silly-Carette, D. Lautru, M. Wong, A. Gati, J. Wiart, and V. F. Hanna, "Variability on the propagation of a plane wave using stochastic collocation methods in a bio electromagnetic application," *IEEE Microwave and Wireless Components Letters*, vol. 19, no. 4, pp. 185-187, 2009.

[14] Z. Zubac, D. D. Zutter, and D. V. Ginste, "Scattering from two-dimensional objects of varying shape combining the method of moments with the stochastic Galerkin method," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 9, pp. 4852-4856, 2014.

[15] P. Manfredi, D. D. Zutter, and D. V. Ginste, "The relationship between Galerkin and collocation methods in statistical transmission line analysis," in *2016 IEEE 25th Conference on Electrical Performance of Electronic Packaging And Systems (EPEPS)*, pp. 35-38, 2016.

[16] D. Poljak, S. Šesnić, M. Cvetković, S. Lallechere, and K. E. K. Drissi, "On some applications of stochastic collocation method in computational electromagnetics: Applications in ground penetrating radar, bioelectromagnetics, grounding systems and buried lines," in *2016 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 1-5, 2016.

[17] G. L. Soares, R. L. S. Adriano, C. A. Maia, L. Jaulin, and J. A. Vasconcelos, "Robust multiobjective TEAM 22 problem: A case study of uncertainties in design optimization," *IEEE Transactions on Magnetics*, vol. 45, no. 3, pp. 1028-1031, 2009.

[18] Z. Ren, C. Park, and C.-S. Koh, "Numerically efficient algorithm for reliability-based robust optimal design of TEAM problem 22," *IEEE Transactions on Magnetics*, vol. 50, no. 2, pp. 661-664, 2014.

[19] A. Pierquin, S. Brisset, T. Henneron, and S. Clénet, "Optimization of the TEAM 22 problem using POD-EIM reduced model," in *2016 IEEE Conference on Electromagnetic Field Computation (CEFC)*, pp. 1-1, 2016.

[20] Z. Wang, W. Che, and L. Zhou, "Uncertainty analysis of the rational function model used in the complex permittivity measurement of biological tissues using PMCT probes within a wide microwave frequency band," *Progress In Electromagnetics Research*, 90, pp. 137-150, 2009.

[21] U. C. Hasar, et al., "Application of a useful uncertainty analysis as a metric tool for assessing the performance of electromagnetic properties retrieval methods of bianisotropic metamaterials," *Progress In Electromagnetics Research*, 128, pp. 365-380, 2012.

[22] B. Jinjun, Z. Gang, and W. Lixin, *Uncertainty Analysis in EMC Simulation Based on Stochastic Collocation Method*. 2015.

[23] L. Teckentrup, P. Jantsch, C. G. Webster, and M. Gunzburger, "A multilevel stochastic collocation method for partial differential equations with random input data," *SIAM/ASA Journal on Uncertainty Quantification*, vol. 3, no. 1, pp. 1046-1074, 2015.

[24] L. Ma, D. Z. Feng, and J. He, "Reduced-dimension clutter suppression method for airborne multiple-input multiple-output radar based on three iterations," *Radar Sonar & Navigation Iet*, vol. 9, no. 3, pp. 249-254, 2015.

Numerical Simulation and Experimental Study of ISAR Imaging of Spherical Convergent Flap Nozzle

Yichao Liang [1], Qingzhen Yang [2], Yongqiang Shi [2], Jin Bai [2], and Qi Lin [1*]

[1] School of Aerospace Engineering
Xiamen University, Xiamen, Fujian 361005, China
247142586@qq.com, *qilin@xmu.edu.cn

[2] School of Power and Energy
Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China

Abstract ─ The research on the electromagnetic scattering characteristics of the aero-engine nozzle can provide an important reference for the electromagnetic stealth design of aircraft. In this paper, the numerical simulation and experimental measurement are used to study the radar cross section (RCS) and inverse synthetic aperture radar (ISAR) image characteristics of spherical convergence flap nozzle (SCFN). Firstly, the feasibility of using the shooting bouncing ray (SBR) method to calculate the scattering characteristics of electrically large and deep cavity is verified. Secondly, the numerical simulation data and experimental data of electromagnetic scattering of SCFN are processed by ISAR imaging algorithm, the numerical and experimental ISAR images of SCFN are obtained. The results show that SBR can effectively calculate the RCS of the electrically large size cavity within ± 30° of the axis direction. The polarization influences the experimental ISAR image. Both the numerical and the experimental ISAR image can accurately reflect the scattering center distribution inside the SCFN, and the numerical ISAR image is superior to the experimental ISAR image in image entropy and equivalent number of looks (ENL), and is partly inferior to the experimental ISAR image in terms of image average gradient. The results prove the validity and accuracy of the numerical ISAR image.

Index Terms ─ Inverse Synthetic Aperture Radar (ISAR), shooting and bouncing ray (SBR), scattering center, spherical convergence flap nozzle (SCFN).

I. INTRODUCTION

For the military aircraft, the aero-engine nozzle produces a significant impact on the backscattering of aircraft. Therefore, whether the electromagnetic scattering of the engine nozzle can be effectively controlled is directly related to the backward electromagnetic stealing capability of the aircraft. As a representative of a new type of vector nozzle, spherical convergent flap nozzle (SCFN) has superior aerodynamic and stealth performance. Due to its simple structure, large vectoring angle and easy integration with the aircraft, SCFN has been included in the integrated high performance turbine engine technology (IHPTET) program by the United States [1].

Inverse synthetic aperture radar (ISAR) has been widely used in the military field, it has mainly been applied to automatic recognition and stealth of targets in recent years. Because of the particularity of the field, there is less information on the ISAR imaging of real military aircraft. Besides, there are few published results in the ISAR imaging of aircraft with the inlet and nozzle, and it is rare to directly study the ISAR imaging of aero-engine nozzle. ISAR imaging of the engine nozzle can analyze the strong scattering source inside the nozzle and guide its radar cross section (RCS) reduction, which would help improve the backward electromagnetic stealth performance of the aircraft and also provide a reference to the aerodynamic design of the engine nozzle.

In 1990, Lin et al. conducted the first electromagnetic scattering test on the S-shape inlet and obtained the RCS and ISAR images of the S-shape inlet [2]. However, due to the underdeveloped computer graphics technology, the ISAR image was very blurred and could not reflect the distribution of the scattering center clearly. In recent years, Guo et al. tested the electromagnetic scattering characteristics of the axisymmetric nozzle and the double S-shape nozzle respectively, and obtained a relatively clear ISAR image [3, 4]. Although measurement in anechoic chamber is an effective and accurate method, the current testing device is difficult to measure full pitch/yaw angles for large targets. The numerical simulation of ISAR imaging of the target can make up for this shortcoming.

To obtain numerical ISAR image, it is necessary to calculate the scattering data of the target over a range

Submitted On: April 16, 2019
Accepted On: April 12, 2020

of frequencies and angles. For electrically large size cavities, the full wave methods such as method of moment (MOM) or finite element method (FEM) consume too much computational resources [5], so high frequency approximation methods such as shooting and bouncing rays (SBR) are the main methods currently used. In recent years, in order to accelerate the calculation speed and accuracy of ISAR images, the traditional SBR has many improved methods, such as the Ray Travel Map (RTM) method in [6]; image-domain ray tube integral equation method proposed in [7, 8]; Jiang et al. used OpenGL to speed up ray tracing in the SBR [9]; Dong et al. added the EEC method to the SBR to correct the contribution of edge diffraction [10]. However, the models studied by the above scholars are closed targets, and the ISAR images lack experimental verification. Sze studied the ISAR image of electrically large zigzag-grooved structures [11], but the model differs greatly from the actual inlet or nozzle; Pienaar et al. performed numerical simulation and experimental tests on the electromagnetic scattering characteristics of the Boe707 model [12, 13]. However, the inlet of the model was closed and could not reflect the scattering characteristics of the cavity. In summary, it is necessary to conduct numerical simulation and experimental verification of ISAR imaging of electrically large cavity.

In this paper, the numerical calculation and experimental test of electromagnetic scattering characteristics of SCFN under different frequencies and different polarization modes are conducted. Firstly, the RCS of SCFN obtained by SBR and measurement are compared, and the effectiveness of SBR in calculating the electromagnetic scattering characteristics of the electrically large and deep cavity with complex internal structure is verified. Secondly, the numerical ISAR image and experimental ISAR image of SCFN are acquired by imaging algorithm and compared through different criteria. Finally, the effects of frequency, polarization mode and experimental procedure on ISAR images are analyzed.

II. SBR METHOD AND VERIFICATION

For an arbitrary tube, the scattered electric field at observation point A (r,θ,φ) is [14]:

$$\boldsymbol{E}(r,\theta,\varphi)=\frac{e^{-jkr}}{r}(\hat{\theta}A_{\theta}+\hat{\varphi}A_{\varphi}). \tag{1}$$

According to the Huygens' principle, A_θ and A_φ can be obtained by the following formula:

$$\begin{bmatrix}A_{\theta}\\A_{\varphi}\end{bmatrix}=\sum_{i=1}^{N}\begin{bmatrix}B_{\theta}\\B_{\varphi}\end{bmatrix}\left(\frac{jk}{2\pi}\right)(\Delta A)_{exit}S(\theta,\varphi)e^{j\hat{k}r_A}, \tag{2}$$

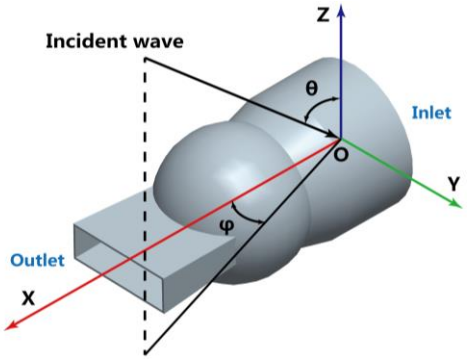
with

$$\begin{aligned}B_{\theta}&=-s_1E_3\cos\varphi-s_2E_3\sin\varphi+s_3(E_1\cos\varphi+E_2\sin\varphi),\\B_{\varphi}&=s_1(E_3\cos\varphi\sin\varphi+\sin\theta E_2)+s_2(-E_1\sin\theta-\\&E_3\cos\theta\cos\varphi)+s_3(E_2\cos\theta\cos\varphi-E_1\cos\theta\sin\varphi).\end{aligned} \tag{3}$$

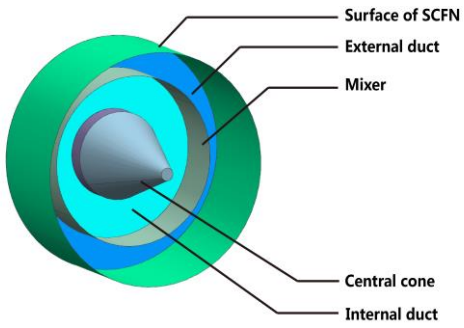
Where N is the total number of the effective ray tubes, r_A is the position vector of the last hit point on the target, $s(\theta,\varphi)$ is the shape function, ΔA_{exit} is the cross section of the exit ray tube. $\boldsymbol{E}(A)=E_1\hat{x}+E_2\hat{y}+E_3\hat{z}$, $\boldsymbol{E}(A)$ is the electric field with each ray tube, $\boldsymbol{s}=s_1\hat{x}+s_2\hat{y}+s_3\hat{z}$, $\boldsymbol{s}$ is the exit ray tube direction, and $\boldsymbol{k}=k_1\hat{x}+k_2\hat{y}+k_3\hat{z}$, $\boldsymbol{k}$ is the observation wave vector.

To improve the calculation accuracy, 1/20 wavelength is chosen as the step size for dividing the ray tube. For accelerating the computational efficiency of the intersection of the ray and the triangle, a data structure based on the Bounding Volume Hierarchy BVH) method is employed.

In order to verify the accuracy of the SBR, the RCS of SCFN calculated by SBR and measurement are compared. The structure of the SCFN is shown in Fig. 1. The incident angle of the plane wave is θ = 90°, φ = -30° to 30°. The SCFN has a central cone and mixer inside, and the inlet of SCFN is disposed as metal terminal. The comparison between the calculated results and the experimental results are shown in Fig. 2.

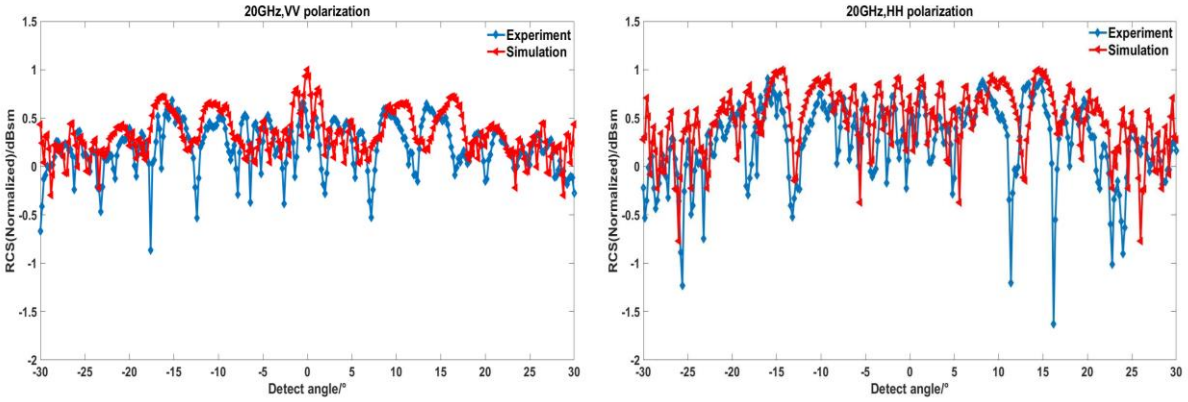


(a) SCFN and Coordinate System



(b) The bottom of SCFN

Fig. 1. SCFN geometric model and local profile.



(a) Vertical polarization (b) Horizontal polarization

Fig. 2. Comparison of SCFN's RCS: "Experiment" (blue dots and line), and "simulation" (red dots and line).

Figure 2 illustrates normalized RCS of SCFN obtained by SBR and measurement at 20 GHz and different polarization modes. It can be seen from Fig. 2 that the RCS of experimental result at 0° is not the peak value because the adjustment of the nozzle central axis is difficult to implement during the test, which causes some deviation in detection angle. From the comparison results of RCS curves in Fig. 2, the simulation results are close to the experimental results. The simulation results of horizontal polarization agree well with the experimental results in the range of ±30° while there are certain errors at vertical polarization. The difference between the numerical results and the experimental results is partly due to the inevitable error in model manufacturing. The anechoic chamber background also has a cluttered effect on the test results, which can be reduced by background vector canceling, but the influence cannot be completely eliminated. In general, the SBR has certain accuracy and reliability when calculating electromagnetic scattering characteristics of electrically large and deep cavity.

III. ISAR IMAGING ALGORITHM

In this paper, the filtered back projection method (FBP) is used to reconstruct the ISAR image. The relationship between the scattering field and the image is given by the following formula [15, 16]:

$$g(x,y)=\int_{\theta_{\min}}^{\theta_{\max}}\int_{k_{\min}}^{k_{\max}} kG(k,\theta)\exp[j2\pi k(y\cos\theta - x\sin\theta)]dkd\theta. \quad (4)$$

Where $g(x,y)$ is the ISAR image of the target, $G(k,\theta)$ is the scattering filed of measured or calculated. $k=2\pi/\lambda$, λ is wavelength, θ is the azimuth angle, as Fig. 3 shows. Then (4) can be decomposed into:

$$p_\theta(l)=\int_0^B (k+k_{\min})G(k+k_{\min},\theta)\exp(j2\pi kl)dk, \quad (5)$$

$$g(x,y)=\int_{\theta_{\min}}^{\theta_{\max}} p_\theta(l)\exp(j2\pi k_{\min}l)d\theta, \quad (6)$$

$$l=y\cos\theta - x\sin\theta. \quad (7)$$

Where $B=k_{\max}-k_{\min}$. Since the frequencies in the calculation and measurement are discrete, it is necessary to discretize (5) and (6). If $k=nB/N$, n=0, 1…, N-1, N is number of sampling frequency points, then (5) can be rewritten as:

$$p_\theta(l)=\sum_{n=0}^{(N-1)B/N}(\frac{B}{N}n+k_{\min})G(n,\theta)\exp(j2\pi\frac{B}{N}nl). \quad (8)$$

The projection line l is equidistantly discretized according l_m, $l_m=m/B$, m=0,1,…,N-1:

$$\text{let } G_\theta(n)=(\frac{B}{N}n+k_{\min})G(n,\theta), \text{ then,}$$

$$p_\theta(l_m)=\sum_{n=0}^{N-1}G_\theta(n)\exp(j2\pi\frac{m}{N}n)=IFFT[G_\theta(n)]. \quad (9)$$

Where $p_\theta(l_m)$ is the projection value obtained from different l_m. The projection line l in the process of integration changes with θ. For $g(x,y)$ at any point, each θ corresponds to a different l. Therefore, $p_\theta(l)$ is the projection value obtained from the discrete l, and $p_\theta(l_m)$ needs to be obtained by interpolating $p_\theta(l_m)$:

$$p_\theta(l)=\text{interp}[p_\theta(l_{m-1}),p_\theta(l_m)]. \quad (10)$$

"Interp" means the interpolation method. In this paper, linear interpolation is selected as interpolation method. Finally, the ISAR image can be obtained by integrating angle:

$$g(x,y)=\int_{\theta_{\min}}^{\theta_{\max}} p_\theta(l)e^{j2\pi k_{\min}l}d\theta=\sum_\theta p_\theta(l)e^{j2\pi k_{\min}l}. \quad (11)$$

It can be seen from the above analysis that the formula derivation process of the FBP method is strict, and only the interpolation process brings errors. In

addition, in order to reduce the side lobes, Kaiser window function is used.

IV. ISAR IMAGING STUDY OF SCFN

A. Numerical simulation of ISAR imaging of SCFN

In this paper, SBR is used to calculate the electromagnetic scattering characteristics of SCFN. The frequency is the same as the experiment, which is 3 to 6 GHz, 8 to 12 GHz, 19 to 25 GHz respectively; the polarization modes are horizontal polarization and vertical polarization; the incident angle is θ = 90°, φ = -20°~20°. The calculation results are given together with the experimental results for comparison.

B. Experimental test of ISAR imaging of SCFN

The measurement was conducted in the National Defense Key Laboratory of UAV Special Technology of Northwestern Polytechnical University. The experiment used a stepped-frequency test system. The stepped-frequency signal is generated by Agilent E8363A VNA. The experiment uses broadband horn antenna HD-1060, its bandwidth is 1-6GHz, the gain is 6-13dB; standard gain horn antenna HD-100 and HD220, bandwidth is 8.20-12.40GHz and 17.6-26.7GHz respectively, and the gain is 15dB. The cross polarization gain of the antenna is less than 30 dB of the main polarization gain.

To test the electromagnetic scattering characteristics inside the SCFN, the low electromagnetic scattering characteristic carrier (hereinafter referred to as a carrier) was designed and fabricated. The outside of the carrier is covered with radar absorbing material. The SCFN is wrapped in a carrier as shown in Fig. 4.

The measurement system construction is shown in Fig. 5. The vector network analyzer transmits the stepped-frequency signal, which is amplified by a power amplifier and then transmitted, finally received by the horn antenna. To minimize the influence of the coupling signals on measurement accuracy, some radar absorbing materials were placed between the two antennas during the measurement.

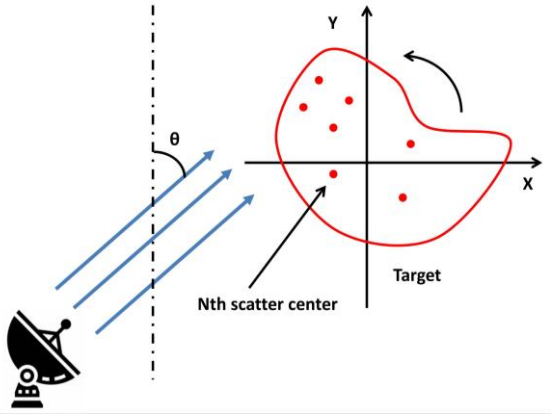
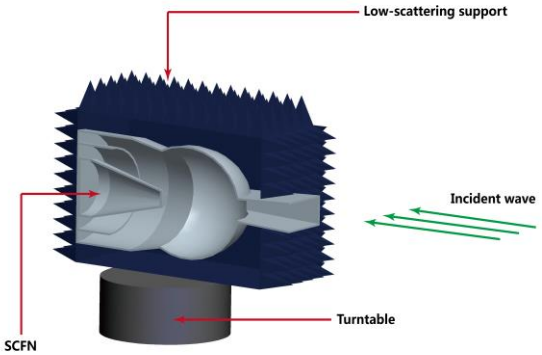
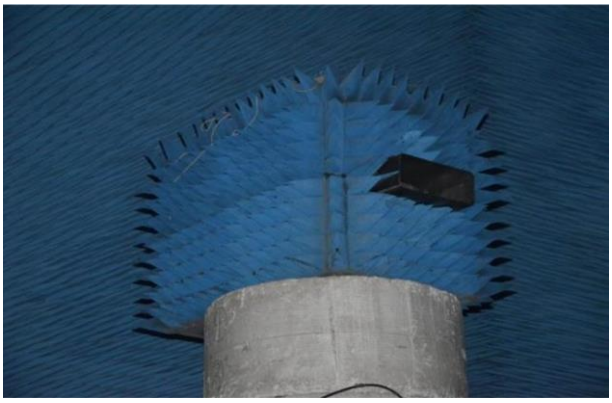


Fig. 3. Schematic diagram of turntable imaging.



(a) The profile of SCFN and carrier



(b) The testing SCFN model

Fig. 4. The structure of testing SCFN.

C. Comparison between numerical simulation and experimental results

The results of numerical simulation and experimental testing of the ISAR imaging of the SCFN are shown in Fig. 6 to Fig. 8. The ISAR image of the cavity structure does not reflect the geometry of the cavity and only gives the distribution of the scattering center inside the cavity. For ease of understanding, the outline of the SCFN is supplemented, as depicted by the red lines in Fig. 6 to Fig. 8. For the sake of simplicity and not to cause misunderstanding, the ISAR image based on numerical simulation is simply referred to as "numerical ISAR image", and the ISAR image based on experimental test is simply referred to as "experimental ISAR image.

Figure 6 shows ISAR imaging results at 3 to 6 GHz. Both the calculation results and the experimental results show that when electromagnetic wave entered inside the SCFN, the bottom of SCFN contributes the most to scattering because the scattering type at the bottom is specular scattering. Since electromagnetic

wave bounces multiple times between the mixer and the center cone, the scattering in this area is also strong. It can be seen from the experimental results in Fig. 6 (b) and Fig. 6 (d) that the carrier also has a certain scattering intensity.

It can be seen from the experimental ISAR images under different polarization modes that the outlet of SCFN shows two spots under vertical polarization and a stripe under horizontal polarization. The experimental results show that polarization has different effects on scattering, mainly because the edge diffraction intensity caused by polarization is different. When the edge is parallel to the direction of polarization, the edge will produce strong edge diffraction; conversely, when the edge is perpendicular to the direction of polarization, the resulting edge diffraction intensity is much weaker. In addition, since the surface of the SCFN is not smooth, electromagnetic waves are diffusely reflected, and the outlet is not completely wrapped by the carrier during measurement, so the outlet can be clearly seen in the experimental results. The outlet cannot be represented in numerical results because all surfaces in the calculation were treated as non-thickness surfaces, and SBR ignores the edge diffraction in the calculation and considers all regions as ideal conductors, diffuse reflection does not exist.

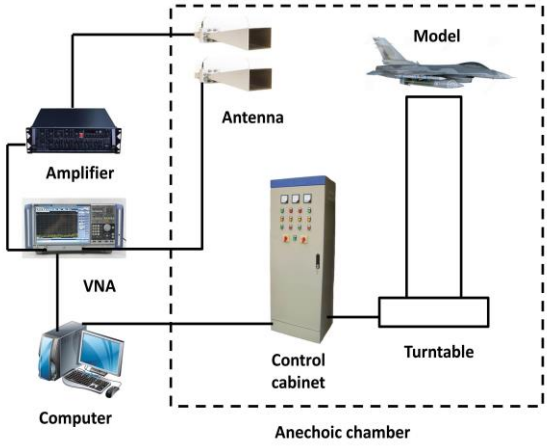
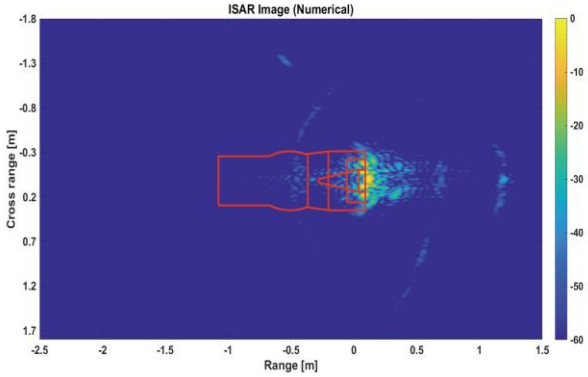
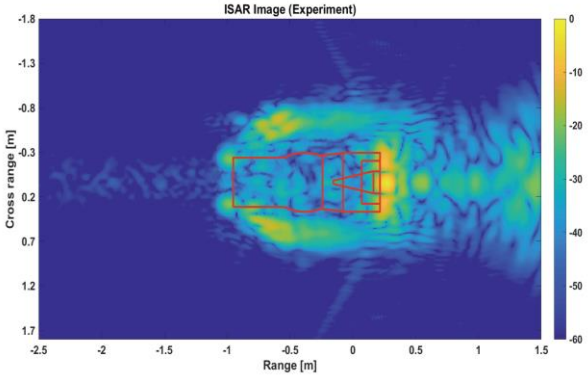


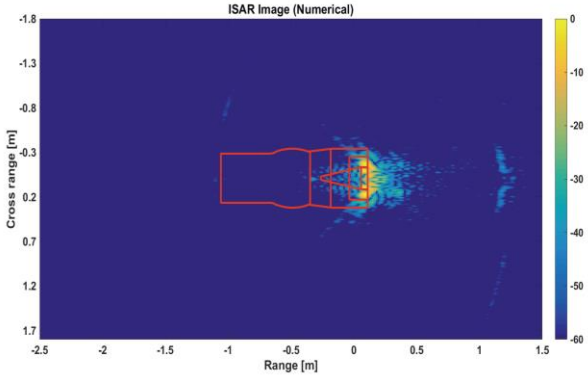
Fig. 5. Diagram of measurement system.



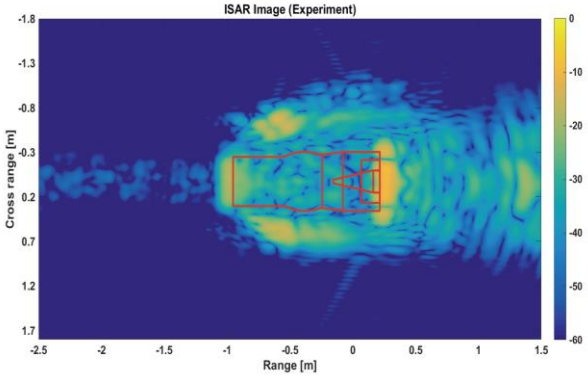
(a) Numerical ISAR Image (vertical polarization)



(b) Experimental ISAR Image (vertical polarization)

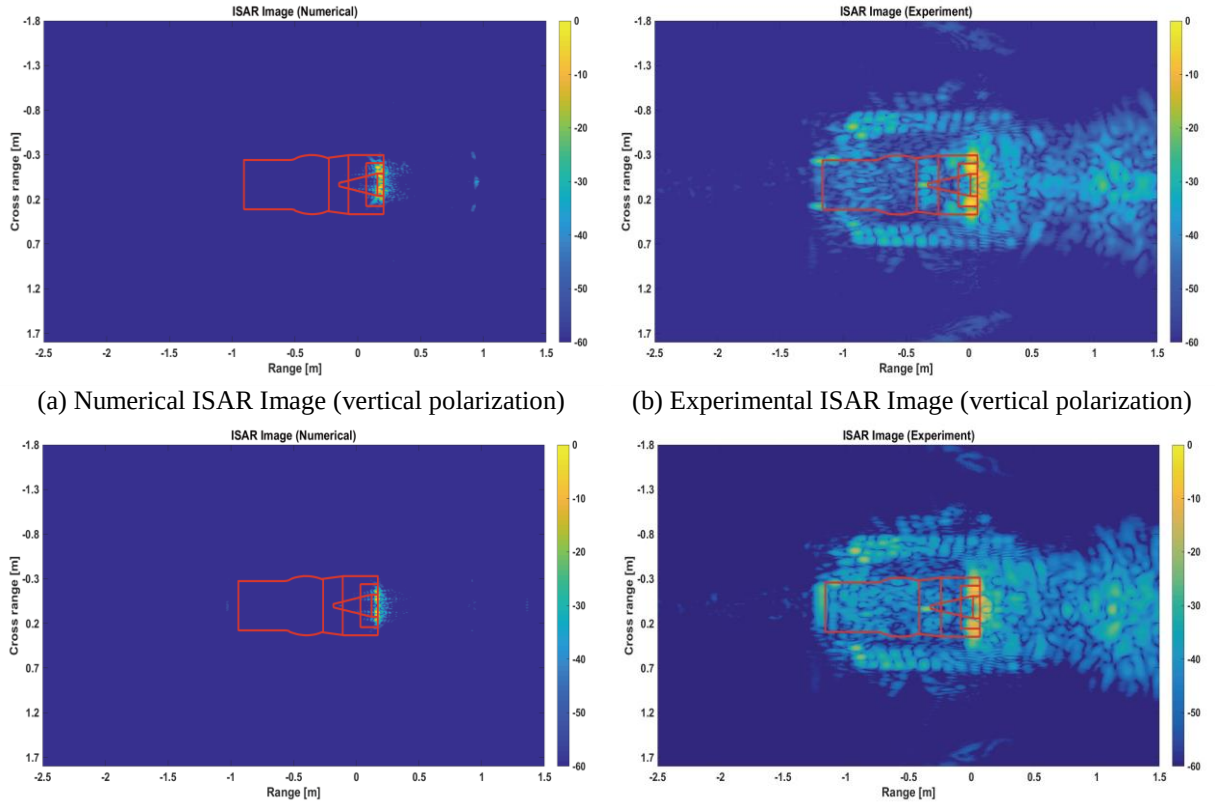


(c) Numerical ISAR Image (horizontal polarization)



(d) Experimental ISAR Image (horizontal polarization)

Fig. 6. Comparison between numerical and experimental ISAR image (3 to 6GHz).



(a) Numerical ISAR Image (vertical polarization)

(b) Experimental ISAR Image (vertical polarization)

(c) Numerical ISAR Image (horizontal polarization)

(d) Experimental ISAR Image (horizontal polarization)

Fig. 7. Comparison between Numerical and Experimental ISAR Image (8 to 12GHz).

Figure 7 shows ISAR imaging results at 8 to 12 GHz. The center cone of SCFN is a truncated cone which has a small flat surface at the top. The type of scattering of this small flat surface is the same as the bottom of SCFN, both of which are specular scattering. However, since the area of the top of center cone is too small, the scattering intensity is much smaller than the bottom of SCFN. In Fig. 7, it is faintly visible that there is a bright spot at the position corresponding to the top of the center cone.

Figure 8 shows the imaging results at 19 to 25 GHz. The scattering intensity of the carrier in Fig. 8 is much smaller than that of Fig. 6 and Fig. 7, because the radar absorbing material has a better absorption effect in the high frequency band and a poor effect in the low frequency band. The bright spots produced by the top of the center cone are clearly visible in Fig. 8. As bandwidth and frequency increase, the resolution of the ISAR increases accordingly, so that the absorbing pyramid on the carrier can be clearly distinguished in Fig. 8. In numerical ISAR images and experimental ISAR images, the region outside SCFN also has scattering intensity, which is mainly due to interference caused by imaging algorithm. In addition, the manufacturing error of SCFN, antenna measurement error, darkroom background interference, etc. will cause disturbance to the experimental ISAR image.

In order to quantitatively evaluate the quality of ISAR images, Shannon entropy, equivalent number of looks (ENL) and image average gradient are taken as the criteria for evaluating the quality of ISAR images [17]. For an ISAR image in size of $M \times N$, define the Shannon entropy *H(f)* as:

$$H(f) = -\sum_{m=1}^{M}\sum_{n=1}^{N} P_{mn} \ln P_{mn}, \tag{12}$$

where

$$P_{mn} = \frac{f(m,n)}{\sum_{m=1}^{M}\sum_{n=1}^{N} f(m,n)}. \tag{13}$$

$f(m,n)$ represents the gray value of the pixel at (m,n) in the ISAR image. The smaller entropy value of ISAR image is, the more information of ISAR image contains, the better the focusing performance of ISAR image has.

ENL is an index to measure the relative intensity of speckle noise in an image. The larger the ENL is, the better the identifiability is, and the weaker the speckle is. It is defined as:

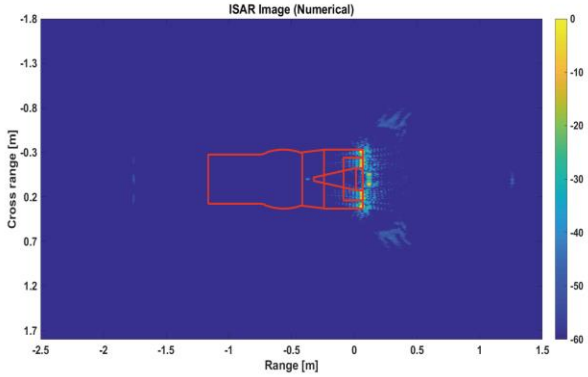
$$ENL = \frac{E(I)^2}{\mathrm{var}(I)}. \quad (14)$$

Where $E(I)$ and $\mathrm{var}(I)$ represent the mean and variance of the image respectively.

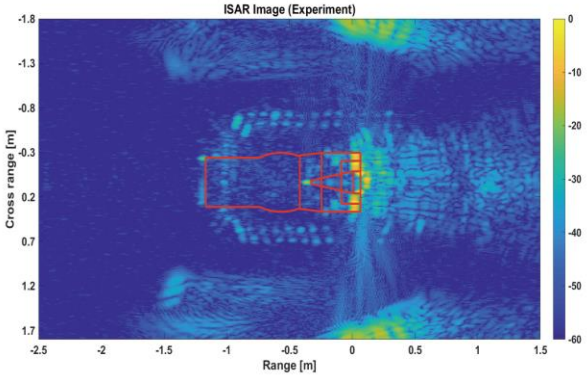
The average gradient of the image indicates the rate at which the image details change and represents the relative clarity of the image. The larger the average gradient is, the clearer the image is. It is defined as:

$$G = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{((\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2) / 2}. \quad (15)$$

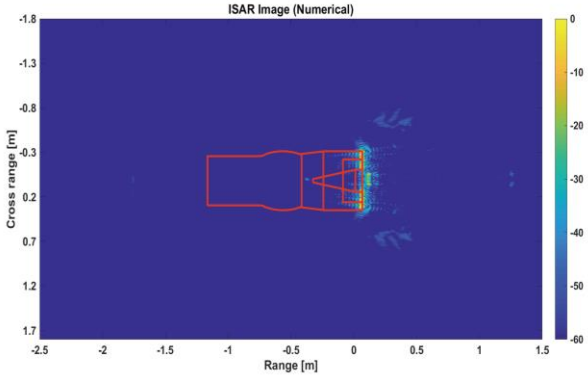
Where $M \times N$ denotes the image size, $\frac{\partial f}{\partial x}$ denotes the horizontal gradient, and $\frac{\partial f}{\partial y}$ denotes the vertical gradient.



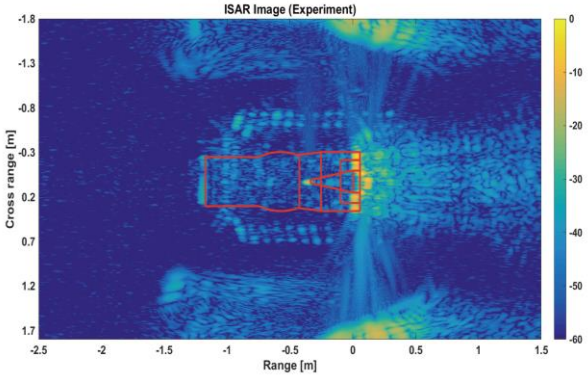
(a) Numerical ISAR Image (vertical polarization)



(b) Experimental ISAR Image (vertical polarization)



(c) Numerical ISAR Image (horizontal polarization)



(d) Experimental ISAR Image (horizontal polarization)

Fig. 8. Comparison between numerical and experimental ISAR image (19~25 GHz).

It can be seen from Table 1 and Table 2 that different polarization modes have little effect on the entropy of ISAR image, and numerical ISAR image entropy is smaller than the entropy of experimental ISAR image, which indicates that the focusing performance of numerical image is better than that of experimental image. The ENL of the numerical ISAR image is larger than that of the experimental ISAR image, indicating that the experimental ISAR image is difficult to identify, which may due to the strong interference in the experimental process. The image average gradient characterizes the intensity of the image change. Since the numerical ISAR image has no background and carrier interference, most areas of the image do not change much, so the image average gradient of the numerical ISAR image is smaller.

Comparing the imaging results we can infer that the area where SCFN contributes mainly to scattering is the part where specular reflection exists, the part where multiple bouncing occurs, and the part where edge diffraction occurs. Since the SBR considers the environment around the target to be vacuum when calculating the target electromagnetic scattering, there is no anechoic chamber background interference, thus the numerical ISAR image is visually clearer than experimental ISAR image. Compared to experimental ISAR images with various disturbances, numerical ISAR images can be better applied to the target recognition. In addition, since the numerical ISAR image can accurately reflect the scattering center distribution inside the SCFN, the numerical ISAR image can be applied to the RCS extraction to judge

the contribution of the internal components of SCFN to total scattering. Taking into account the enormous resources spent on ISAR imaging experiment, the advantages of numerical ISAR images in terms of cost and efficiency are more obvious.

Table 1: Results of each criterion under horizontal polarization

Frequency (GHz)	**3-6**		**8-12**		**19-25**	
Criterion	Num.	Exp.	Num.	Exp.	Num.	Exp.
Image entropy	0.4361	2.9995	0.1123	2.5781	0.2469	3.7061
ENL	21.1856	5.2411	82.7427	7.2330	71.8607	8.1126
Average gradient	0.0076	0.0023	0.0056	0.0097	0.0023	0.0096

*Num.: Numerical; Exp.: Experimental

Table 2: Results of each criterion under vertical polarization

Frequency (GHz)	**3-6**		**8-12**		**19-25**	
Criterion	Num.	Exp.	Num.	Exp.	Num.	Exp.
Image entropy	0.5013	3.0446	0.1968	2.6167	0.3072	3.7858
ENL	20.6227	5.6650	31.7412	7.2334	27.8587	9.1920
Average gradient	0.0107	0.0120	0.0105	0.0122	0.0109	0.0108

*Num.: Numerical; Exp.: Experimental

CONCLUSION

In this paper, the RCS characteristics and ISAR image of SCFN are numerically calculated and experimentally tested. The conclusions are as follows.

(1) The comparison with the existing experimental data shows that the SBR can accurately calculate the scattering characteristics of the electrically large and deep cavity in the axial range of ± 30°.

(2) The experimental results indicate that there are differences in the effects of polarization modes on the scattering characteristics of binary vector nozzles. The main reason for the differences is the relative positional relationship between the surface of the outlet and the polarization direction.

(3) The numerical ISAR image is superior to the experimental ISAR image in image entropy and ENL, and the average gradient of the experimental ISAR image is partly better than the experimental ISAR image. The comparison between the experimental ISAR image and the numerical ISAR image demonstrates that the ISAR image can accurately reflect the scattering center distribution inside the SCFN, so the RCS of the target component can be extracted from the ISAR image, which is also the direction of future work.

AKNOWLEDGEMENT

The author would like to sincerely appreciate the Key Laboratory of UAV Special Technology of Northwestern Polytechnical University for providing technical support for experimental testing, as well as the vigorous support of A. P Chen Weijun and A. P Hu Chufeng during experimental testing and the valuable guidance given in the writing.

REFERENCES

[1] X. Guo, Q. Z. Yang, Y. Q. Shi, and H. Q. Yang, "Effects of vector angle on RCS of spherical convergent flap nozzle," *Journal of Propulsion Technology*, vol. 39, no. 4, pp. 785-791, Apr. 2018.

[2] L. Lin and W. Chen, "Experimental study on radar scattering characteristics of S-shaped inlet," *Aerodynamic Missile Journal*, no. 05, pp. 41-46, May 1990.

[3] X. Guo, Q. Z. Yang, Y. Q. Shi, and H. C. Yang, "RCS measurement of axisymmetric nozzle under different radar overlook angle," *Journal of Propulsion Technology*, vol. 38, no. 8, pp. 1747-1753, Aug. 2017.

[4] X. Guo, Q. Yang, H. Yang, and K. Du, "Radar cross section measurement of double S-shaped nozzle with radar absorbing material," in *Selected Papers of the Chinese Society for Optical Engineering Conferences held October and November 2016*, Ed., pp. 102553H, International Society for Optics and Photonics, 2017.

[5] K. Zhang, C. Wang, and J. Jin, "Broadband monostatic RCS and ISAR computation of large and deep open cavities," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 8, pp. 4180-4193, Aug. 2018.

[6] Y. H. Lee, J. H. Bae, B. S. Kang, S. J. Lee, and K. T. Kim, "Acceleration of monostatic shooting and bouncing ray method for inverse synthetic aperture radar image generation," *Microwave and Optical Technology Letters*, vol. 58, no. 8, pp. 1848-1855, Aug. 2016.

[7] X. Y. He, X. Y. Zhou, and T. J. Cui, "Fast 3D-ISAR image simulation of targets at arbitrary aspect angles through nonuniform fast fourier transform (NUFFT)," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 5, pp. 2597-2602, May 2012.

[8] J. Lee, D. Yun, H. Kim, W. Yang, and N. Myung, "Fast ISAR image formations over multiaspect angles using the shooting and bouncing rays," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 6, pp. 1020-1023, June 2018.

[9] W. Jiang, M. Zhang, D. Nie, and Y. Jiao, "Improved GO/PO method and its application to wideband SAR image of conducting objects over rough surface," *Waves in Random and Complex Media*, vol. 28, no. 2, pp. 310-325, Feb. 2018.

[10] C. Dong, H. Ren, H. Yin, and C. Wang, "Efficient

full-polarized ISAR image simulation of complex target with rough surfaces," *Acta Aeronautica et Astronautica Sinica*, vol. 37, no. 4, pp. 1272-1280, Apr. 2016.

[11] K. Y. Sze and S. Kashyap, "Scattering center modeling of electrically large zigzag-grooved structures," Ed., pp. 6401-6405, IEEE, 2012.

[12] C. Pienaar, J. W. Odendaal, J. C. Smit, J. E. Cilliers, and J. Joubert, "RCS results for an electrically large realistic model airframe," *Applied Computational Electromagnetics Society Journal*, vol. 33, no. 1, pp. 87-90, Jan. 2018.

[13] J. Joubert, C. Pienaar, J. C. Smit, J. W. Odendaal, and J. E. Cilliers, "RCS validation of asymptotic techniques using measured data of an electrically large complex model airframe," *Applied Computational Electromagnetics Society Journal*, vol. 32, no. 1, pp. 60-67, Jan. 2017.

[14] H. Ling, R. C. Chou, and S. W. Lee, "Shooting and bouncing rays: calculating the RCS of an arbitrarily shaped cavity," *Antennas & Propagation IEEE Transactions on*, vol. 37, no. 2, pp. 194-205, Feb. 1989.

[15] C. Hu, N. Li, W. Chen, and L. Zhang, "High-precision RCS measurement of aircraft's weak scattering source," *Chinese Journal of Aeronautics*, vol. 29, no. 3, pp. 772-778, May 2016.

[16] N. Li, Z. Xu, C. Hu, J. Dang, and S. Guo, "Research on accurate radar cross section measurement method based on imaging extraction," *Chinese Journal of Scientific Instrument*, vol. 38, no. 1, pp. 74-82, Jan. 2017.

[17] J. Y. Wei and Z. Yan, "Research on ISAR image quality evaluation," *Systems Engineering and Electronics*, vol. 37, no. 2, pp. 298-303, Feb. 2015.

A Source Signal Recovery Method for Underdetermined Blind Source Separation based on Shortest Path

Chuanchuan Wang and Rui Jia

State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System
Luoyang, 471003, China
wangchuan1083@126.com, jiarui315@163.com

Abstract ─ A new shortest path source recovery algorithm is presented for source signal recovery issue in underdetermined blind source separation, by which the source signals can be recovered in case the observed signals are no less than two dimensions. In this algorithm, two adjacent observed signals are taken everytime among *m* observed signals, marked as the i^{th} and j^{th} signals, and form a two-dimensional observed signal combination, $i=1,2,\ldots,m-1, j=i+1$. The first and m^{th} signals are used to form another two-dimensional observed signal combination, then *m* two-dimensional observed signal combinations are obtained. The number of source signals is *n*, the *n* signals can be obtained respectively after signal recovery by using each two-dimensional observed signal combination A matrix (i.e., $\tilde{\boldsymbol{s}}(p,q,:)=\left[\hat{\boldsymbol{s}}(1,1,:),\hat{\boldsymbol{s}}(1,2,:),\cdots,\hat{\boldsymbol{s}}(1,n,:),\hat{\boldsymbol{s}}(2,1,:),\cdots,\hat{\boldsymbol{s}}(m,n,:)\right]^{\mathrm{T}}$) which is a *mn*-dimensional vector combination matrix can be obtained using each signal combination A *mn*×*mn*-dimensional square matrix can be gotten by calculating the vector angle between rows of $\tilde{\boldsymbol{s}}(p,q,:)$ matrix, for the first *n* rows×*mn* columns, position the vector where the matrix elements are between 0 and θ_0 . The mean is calculated for the signal vector with angle smaller than θ_0 as the estimate of the source signals. Thus, the estimates $\hat{\boldsymbol{s}}(1,:),\hat{\boldsymbol{s}}(2,:),\cdots,\hat{\boldsymbol{s}}(n,:)$ of *n* source signals can be obtained eventually. The method presented provides a new option for solving underdetermined blind source signal recovery problem.

Index Terms ─ Shortest Path Method, Source Signal Recovery, Sparsity, Underdetermined Blind Source Separation.

I. INTRODUCTION

Underdetermined blind source separation (UBSS) is a kind of signal processing techniques to estimate the source signals only using observed signals when the source signal prior information and propagation channel parameters are unknown and the number of the observed signals is smaller than the source signals. Recently underdetermined blind source separation has become a research hotspot in international signal processing community, which is generally applicable to the fields of biomedicine, mobile communication, radar signal processing, underwater acoustic signal processing, image processing, voice signal processing, and mechanical fault diagnosis, etc. For example, in the field of communication, it can be used for code division multiple access multi-user detection, interference suppression, noise cancellation, etc. to improve the call quality; in the field of radar signal processing, it can be used for distributed radar interference suppression, signal sorting, etc. to improve echo Signal anti-interference ability, realize radar, communication signal sorting, etc. under the condition of lack of prior information, spectrum aliasing, and on the same frequency [1,2]. But it is also a difficulty issue due to the number of the observed signals is always smaller than the source signals. The way for solving this problem generally divided into two steps: estimate the mixing matrix using observed signals, then recovery the source signals employing the mixing matrix estimated and observed signals [1,2]. The inverse matrix cannot be solved directly to recovery source signals since the mixing matrix is underdetermined. Therefore, the recovery of source signal is a complex issue. The recovery result of the source signals directly relates to the signal blind separation processing. In short, source signal recovery algorithm research is of important theoretical and practical significance.

In [3-5], it is assumed the source signals have strict orthogonality or quasi-orthogonality in time-frequency domain, which means they are not coincident or almost not coincident at some points, and the source signal separation is realized by time-frequency mask. Degenerate unmixing estimation technique (DUET) presented by Yilmaz et al. [3] is a typical time-frequency masking method. The DUET proposed by Cobos et al. [4] improves the accuracy in signal separation. However, both the methods presented are only applicable to the situation when the observed

Submitted On: October 29, 2019
Accepted On: March 1, 2020

signals are two dimensional. In [5], the hypothesis of orthogonality is relaxed, and it is only required that the number of source signals that exist concurrently at the time-frequency point is smaller than observed signals. The assumptions are further relaxed in [6], in which it is only required that the number of the source signals existed concurrently at the time-frequency point is no more than the sensors. The main problem of these methods is the requirements on strict orthogonality or quasi-orthogonality of signals in time-frequency domain are still too rigid.

Bofill et al. [7] presented a source signal sparsity-oriented shortest path source signal recovery algorithm which is simple and efficacious, yet it is only applicable to two-dimensional observed signal. Georgiev et al. [8] proved that, l_1 norm minimizing method is equivalent to the shortest path method. The main problem of the shortest path is only applicable to the two-dimensional observed signals.

Xiao et al. [9] presented a statistically sparse decomposition principle (SSDP). The source signal is estimated using correlation coefficient of minimized source signal within a fixed time interval in the algorithm. It is required the number of non-zero source signals within the interval no more than two, thence it is not applicable to the recovery of underdetermined source signals that are not sufficiently sparse. Zhao et al. [10] expanded SSDP algorithm and obtained the Statistically Non-Sparse Decomposition Principle aimed at two-dimensional observed signals.

Compressed sensing theory is also used for recovery of underdetermined blind source signal. When the estimate of the mixing matrix has been completed, the recovery of underdetermined blind source signal is similar to compressed sensing reconstruction model. Fu et al. compared the three methods (greedy algorithm, l_1 norm algorithm and smooth l_0 norm algorithm) and proposed SCMP algorithm and plane pursuit algorithm [11,12] which are more accurate and less time waste. In [13], a new algorithm based on artificial neural network was introduced. Compressed sensing theory is applied, with the main problem of requiring the source signals have high sparsity, large data sampling points, and high calculation load.

In this paper, we propose a modified shortest path algorithm for UBSS problem. By employing the proposed algorithm, underdetermined blind source signal recover can be realized when the antenna array is more than or equal to two dimensions. The proposed algorithm provides a new technical approach to solve the difficulty of source signal recovery for UBSS.

The remainder of the paper is organized as follows. Section II presents the model for underdetermined blind source separation problem. Section III gives the new source signal recovery method with the modified shortest path algorithm. Section IV describes simulation results that illustrates the effectiveness of the proposed method. Finally, the conclusions are drawn in Section V.

II. MODEL FOR UNDERDETERMINED BLIND SOURCE SEPARATION

The general goal of blind source separation is to recover source signals from observed signals. The aliasing of the source signals to observed signals may be linear instantaneous, convolutional or nonlinear. Research is carried out for linear instantaneous aliasing of source signals in this paper. It is assumed that, there are n source signals expressed as $\boldsymbol{s}(t)=[s_1(t),s_2(t),...,s_n(t)]^{\mathrm{T}}$, where the superscript T means transposition and it has the same meaning in the following. The number of signal sampling points is $t=1,2,\cdots,T_0$. Signal aliasing will occur during the propagation of n source signals and the reception by sensors. The m observed signals received by antenna array are expressed as $\boldsymbol{x}(t)=[x_1(t),x_2(t),...,x_m(t)]^{\mathrm{T}}$. Any observed signal $\boldsymbol{x}(t)$ is the aliasing of the source signals $\boldsymbol{s}(t)$. The mathematical model of linear instantaneous aliasing blind source separation is:

$$\boldsymbol{x}(t)=\boldsymbol{A}\boldsymbol{s}(t)+\boldsymbol{n}(t). \tag{1}$$

Where, the prior information of mixing matrix $\boldsymbol{A}$ and source signals $\boldsymbol{s}(t)$ cannot be measured, only $\boldsymbol{x}(t)$ can be measured. $\boldsymbol{n}(t)$ refers to the additive noise aliased in the process of signal reception.

When the number of observed signals m is smaller than the number of source signals n, the blind source separation is referred to as underdetermined blind source separation. The mixing matrix $\boldsymbol{A}$ is column dissatisfaction rank in underdetermined blind source separation, so $\boldsymbol{s}(t)$ cannot be obtained by the inverse matrix of $\boldsymbol{A}$. Therefore, source signal recovery issue is a huge challenge in underdetermined blind source separation.

III. NEW SOURCE SIGNAL RECOVERY METHOD WITH THE SHORTEST PATH ALGORITHM

A. Sparsity-based underdetermined blind source signal recovery model

Blind source separation based on sparse representation can be used to solve the problem of underdetermined blind source separation. For blind source separation of sparse signals, it can be represented as below to seek for optimal solution:

$$\min_{\boldsymbol{A},\boldsymbol{s}} \frac{1}{2\sigma^2}\left\|\boldsymbol{A}\boldsymbol{s}-\boldsymbol{x}\right\|^2+\sum_{i,t}\left|\boldsymbol{s}_i(t)\right|. \tag{2}$$

Where, σ^2 is the noise variance, $\left\|\boldsymbol{A}\boldsymbol{s}-\boldsymbol{x}\right\|^2$ is the sum

of square of reconstructed error. The final item is the non-sparse penalty (assuming the source signals are independent of each other). Many variables require optimization, and the model can be simplified as the following when the mixing matrix $\boldsymbol{A}$ is known:

$$\min_{s(t)} \frac{1}{2\sigma^2} \left\| \boldsymbol{A}\boldsymbol{s}(t) - \boldsymbol{x}(t) \right\|^2 + \sum_i^n \left| \boldsymbol{s}_i(t) \right|, \quad t = 1, 2, \mathrm{L}\ , T_0. \quad (3)$$

In the absence of noise, equation (3) can be simplified as:

$$\min_{\boldsymbol{s}(t)} \sum_i^n \left| \boldsymbol{s}_i(t) \right|, \quad t = 1, 2, \mathrm{L}\ , T_0. \quad (4)$$

The total optimization for underdetermined source signal recovery can be broken down into T_0 sub-optimization items.

B. Source signal recovery algorithm with the shortest path method

The shortest path method is a simple and effective source signal recovery algorithm for underdetermined blind source separation. It is suitable for the recovery of the source signal which is sufficiently sparse and the observed signals is two-dimensional. When the source signals are not sufficiently sparse in the time domain, they can be sparsely represented adopting time-frequency transformation or wavelet transformation method, therefore the source signal recovery can then be realized adopting the shortest path method. The source signal recovery issue can be converted into the solving of the optimization described below according to sparse component analysis theory:

$$\begin{cases} \min\limits_{\boldsymbol{s}(t)} \sum\limits_{i=1}^{n} | \boldsymbol{s}_i(t) |, \\ \boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) = \sum\limits_{i=1}^{n} \boldsymbol{a}_i \boldsymbol{s}_i(t). \end{cases} \quad (5)$$

Where, $\boldsymbol{x}(t)$ is the observed signals and the number is m; $\boldsymbol{A}$ is the mixing matrix; the number of source signals is n; $\boldsymbol{a}_i$ is the i^{th} column of the mixing matrix and $\boldsymbol{s}_i(t)$ is the i^{th} source signal. To minimize $\sum_{i=1}^{n} \boldsymbol{s}_i(t)$ at the time is to carry out linear decomposition for the observed signals in the directions of two columns of the mixing matrix and to find out the shortest path from the origin to the observed signals. When there are two observed signals, the method to solve the question is shown in Fig. 1. For minimizing $\sum_{i=1}^{n} | \boldsymbol{s}_i(t) |$ at the time, it can be known from Fig. 1, the shortest path from the origin to observed signal $\boldsymbol{x}$ is the two vectors ($\boldsymbol{a}$ and $\boldsymbol{b}$), which are closest to the angle of $\boldsymbol{x}$.

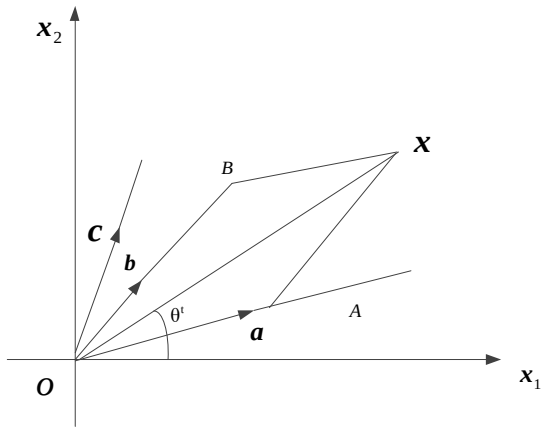


Fig. 1. Schematic for the shortest path method for the two-dimensional observed signals.

For the situation when there are more than two observed signals, the original shortest path method is improved in this paper, therefore it is applicable to the recovery of the source signals with two or more observed signals. The method is taking two adjacent observed signals every single time among m observed signals and express them as the i^{th} and j^{th} observed signals, where, $i = 1, 2, \cdots, m-1$, $j = i+1$, which means the observed signals processed each time is the combination of two adjacent observed signals, then $m-1$ combinations are obtained. The first and the m^{th} observed signals are taken to form another two-dimensional observed signal combination. Then m two-dimensional observed signal combinations can be obtained. For each combination, corresponding source signals can be recovered adopting the original shortest path method, then there are m groups of source signals recovered. As the number of source signals is n, then n signals can be obtained respectively by using each of the m two-dimensional observed signal combinations obtained from the original observed signals. Assuming that the signals obtained from separation of each combination is expressed as $\hat{\boldsymbol{s}}(i, k, :)$, where, $i = 1, 2, \cdots, m$ refers to the serial number of each two-dimensional observed signal combination; $k = 1, 2, \cdots, n$ is the serial number of the signal obtained from separation of each two-dimensional combination, : refers to the number of sampling points. A new matrix is gotten using the signals obtained from the separation of the m groups, and it can be expressed as:

$$\tilde{\boldsymbol{s}}(p, q, :) = [\hat{\boldsymbol{s}}(1,1,:), \hat{\boldsymbol{s}}(1,2,:), \cdots, \hat{\boldsymbol{s}}(1,n,:), \hat{\boldsymbol{s}}(2,1,:), \cdots, \hat{\boldsymbol{s}}(m,n,:)]^{\mathrm{T}}. \quad (6)$$

$\tilde{\boldsymbol{s}}(p, q, :)$ is a $mn \times T_0$ dimensional vector combination

matrix. *mn* refers to the number of the separated signals. T_0 refers to the number of sampling points of the signal.

For matrix $\tilde{\boldsymbol{s}}(p,q,:)$, calculating the vector angles between its rows, then a *mn*×*mn*-dimensional square matrix *Q* can be obtained. For the first *n* rows×*mn* columns of the square matrix, check if the matrix element is between 0 and threshold value θ_0. If yes, it means that the intersection angle of the signals in matrix $\tilde{\boldsymbol{s}}(p,q,:)$ is smaller than θ_0, which reflects the strong similarity of the signals. The mean of the signal with angles less than θ_0 can be calculated respectively as the estimate of the source signals. Thus, the estimates $\hat{s}(1,:),\hat{s}(2,:),\cdots,\hat{s}(n,:)$ of *n* source signals (the source signals to be recovered) are gotten eventually.

The threshold θ_0 is a critical parameter in this algorithm. When determining θ_0, it should consider if the mixing matric can be estimated accurately. In this paper the threshold is determined based on the research of Zayyani et al. [14] and Cramér–Rao bound of mixing matrix. The threshold should be larger than the least angle that the mixing matrix can be estimated correctly. The extensive simulations show that the algorithm proposed in this paper has good robustness when the threshold θ_0 is larger. The simulation shows that, when the value of θ_0 is within $[15^\circ, 35^\circ]$, the difference of source signals recovery result is not obvious. A larger θ_0 is more favourable for signal recovery results since more signals can be integrated to obtain the estimates of source signals.

Specific steps of the algorithm presented in this paper are shown as follow:

Step 1: Pre-processing is carried out for the observed signals $\boldsymbol{x}(t)$, which can remove the column vectors with all components are zero. Then unification of direction is conducted.

Step 2: In *m* observed signals $\boldsymbol{x}(t)$ obtained from one measurement, $\boldsymbol{x}(t)=[\boldsymbol{x}_1(t),\boldsymbol{x}_2(t),\cdots,\boldsymbol{x}_m(t)]^{\mathrm{T}}$, the sampling time $t=1,2,\cdots,T_0$. Two adjacent observed signals $\boldsymbol{x}_i(t)$ and $\boldsymbol{x}_j(t)$ are selected each time, $i=1,2,\cdots,m-1$, $j=i+1$. The first and the *mth* observed signals are taken to form a two-dimensional observed signal combination. Then m two-dimensional observed signal combinations can be obtained in total, which are expressed as $\boldsymbol{x}_k(t)=[\boldsymbol{x}_i(t),\boldsymbol{x}_j(t)]^{\mathrm{T}}$ respectively, where, $k=1,2,\cdots,m$;

Step 3: Calculate the angle of each base vector of mixing matrix $\boldsymbol{A}$: the angle of base vector is defined as $\theta^{A_j}=\tan^{-1}(\boldsymbol{A}_j^2/\boldsymbol{A}_j^1)$, $\boldsymbol{A}_j$ refers to the j^{th} column vector of the mixing matrix, $j=1,2,\cdots,n$; *n* is the number of source signals, the superscripts 2 and 1 mean the second and the first row of the column vector respectively;

Step 4: The angle of the observed signal vector $\boldsymbol{x}^t$ in each combination is calculated respectively adopting the original shortest path algorithm for the combination of m two-dimensional observed signals at each observed moment;

Step 5: Find out the two base vector angles closest to the observed signal vector angle θ^t at the moment, and record the two row vectors $\boldsymbol{a}^i$ and $\boldsymbol{b}^i$ in corresponding mixing matrix, where, $\boldsymbol{a}^i,\boldsymbol{b}^i\in\boldsymbol{A}$ and *i* refers to the serial number of the i^{th} combination in the *m* combinations, $i=1,2,\cdots,m$;

Step 6: Assume $\boldsymbol{A}_r=[\boldsymbol{a}^i\ \ \boldsymbol{b}^i]$, $\boldsymbol{A}_r$ refers to a submatrix of 2×2 formed by two rows of $\boldsymbol{a}^i$ and $\boldsymbol{b}^i$ in mixing matrix $\boldsymbol{A}$. $\boldsymbol{a}^i$ and $\boldsymbol{b}^i$ are the two vectors closest to $\boldsymbol{x}^t$ at the moment of *t*, then make $\boldsymbol{W}_r=\boldsymbol{A}_r^{-1}$;

Step 7: The source signals at sampling moment *t* are recovered using the following equation:

$$\begin{aligned}&\boldsymbol{s}_r^t=\boldsymbol{W}_r\boldsymbol{x}^t,\\&\boldsymbol{s}_j^t=0\quad for\quad j\neq\boldsymbol{b}^i,\boldsymbol{a}^i.\end{aligned}\tag{7}$$

Where, $\boldsymbol{s}_r^t$ is the component of $\boldsymbol{x}$ in the directions of vector $\boldsymbol{a}^i$ and $\boldsymbol{b}^i$;

Step 8: For each two-dimensional observed signal combination $\boldsymbol{x}_k(t)=[\boldsymbol{x}_i(t),\boldsymbol{x}_j(t)]^{\mathrm{T}}$, $k=1,2,\cdots,m$, corresponding source signals are recovered adopting the original shortest path method, and there are m groups of source signals recovered. As the number of source signals is n, so n signals can be obtained after recovery by using each of m two-dimensional observed signal combinations. Assuming the signal obtained from separation of each combination is expressed as $\hat{\boldsymbol{s}}(i,k,:)$, where, $i=1,2,\cdots,m$ refers to each two-dimensional observed signal combination; $k=1,2,\cdots,n$ refers to the signal obtained from separation of each two-dimensional combination; : means the number of sampling points;

Step 9: A new matrix is formed using the signals obtained from the separation of *m* two-dimensional observed signal combinations, which can be expressed as $\tilde{\boldsymbol{s}}(p,q,:)=[\hat{\boldsymbol{s}}(1,1,:),\hat{\boldsymbol{s}}(1,2,:),\cdots,\hat{\boldsymbol{s}}(1,n,:),\hat{\boldsymbol{s}}(2,1,:),\cdots,\hat{\boldsymbol{s}}(m,n,:)]^{\mathrm{T}}$. $\tilde{\boldsymbol{s}}(p,q,:)$ is a $mn\times T_0$-dimensional vector combination matrix.

Step 10: A *mn*×*mn*-dimensional square matrix $\boldsymbol{Q}$ can be obtained by finding the angle between the row vectors of matrix $\tilde{\boldsymbol{s}}(p,q,:)$. For the first n rows×*mn* columns of the square matrix, check if the matrix element is larger than 0 but smaller than θ_0, meaning

that the angle of the single vector in matrix $\tilde{\boldsymbol{s}}(p,q,:)$ is smaller than θ_0. The mean is calculated for the signal vectors with angle smaller than θ_0 as the estimate of the source signal. Thus, the estimates $\hat{\boldsymbol{s}}(1,:),\hat{\boldsymbol{s}}(2,:),\cdots,\hat{\boldsymbol{s}}(n,:)$ of n source signals (i.e., the source signals to be recovered) can be obtained eventually.

IV. SIMULATION AND ANALYSIS

The simulation platform is a DELL9020MT computer, Intel(R) Core(TM) i7-4770 CPU @3.40GHz, 64-bit Windows operating system. MATLAB software is used for the simulation experiment. The signal-to-noise ratio of the observed signals changes from 8 to 20dB and Monte Carlo simulation at each signal-to-noise ratio is carried out for 500 times. The performance indexes for evaluating the recovery effect of the source signal include the average separated signal to interference ratio $\overline{\mathrm{SIR}} = mean(\mathrm{SIR}_i)$ and the average similarity coefficient $\bar{\xi} = mean(\xi_{ij})$, $i, j = 1,2,\cdots,n$. Where, SIR and ξ are the separated signal to interference ratio and similarity coefficient respectively. The formulas for calculation of them are shown in (8) and (9), where, $\hat{\boldsymbol{s}}_i(t)$ and $\hat{\boldsymbol{s}}_j(t)$ are the i^{th} and j^{th} signals respectively. $\boldsymbol{s}_i(t)$ refers to the i^{th} source signal:

$$\mathrm{SIR}_i = 10\lg\left(\frac{\sum_{t=1}^{T_0}\boldsymbol{s}_i^2(t)}{\left[\hat{\boldsymbol{s}}_i(t)-\boldsymbol{s}_i(t)\right]^2}\right), \tag{8}$$

$$\xi_{ij} = \frac{\left|\sum_{t=1}^{T_0}\boldsymbol{s}_i(t)\hat{\boldsymbol{s}}_j(t)\right|}{\sqrt{\sum_{t=1}^{T_0}\boldsymbol{s}_i^2(t)\sum_{t=1}^{T_0}\hat{\boldsymbol{s}}_j^2(t)}}. \tag{9}$$

The threshold value θ_0 for signal integration is set at 20°.

Experiment 1:

There are 5 source signals. Assuming the source signals are sufficiently sparse in time domain, then the signal types and parameter settings are as follows:

s_1 is a conventional pulse signal, carrier frequency fc_1=5MHz, pulse width $t_{r1}=10\mu s$, pulse repetition period $T_{r1}=100\mu s$, pulse start time $t_{01}=0$;

s_2 is a conventional pulse signal, carrier frequency fc_2=5MHz, pulse width $t_{r2}=7\mu s$, pulse repetition period $T_{r2}=100\mu s$, pulse start time $t_{02}=10\mu s$;

s_3 is a linear FM signal, carrier frequency fc_3=5MHz, pulse width $t_{r3}=10\mu s$, pulse repetition period $T_{r3}=100\mu s$, pulse start time $t_{03}=20\mu s$, instantaneous bandwidth is $B_3=10MHz$;

s_4 is a linear FM signal, carrier frequency fc_4=5MHz, pulse width $t_{r4}=8\mu s$, pulse repetition period $T_{r4}=100\mu s$, pulse start time t_{04}=30μs, intrapulse bandwidth $B_4=15MHz$;

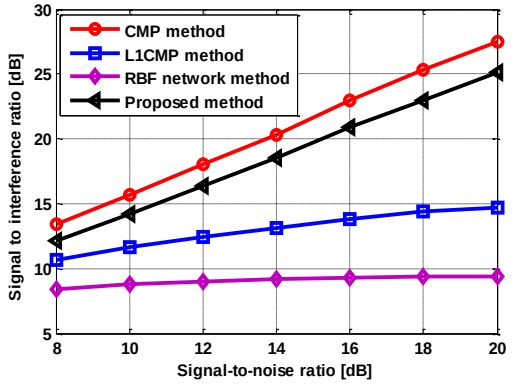
s_5 is a sinusoidal phase-modulated signal, carrier frequency $f_{c5}=5MHz$, pulse width $t_{r5}=8\mu s$, pulse repetition period $T_{r5}=100\mu s$, pulse start time $t_{05}=40\mu s$, modulated signal frequency $f_{a5}=100kHz$, modulation index $a_5=5$.

The sampling frequency of the receiver is 50MHz; the number of sampling points is 10,000 and the dimensional of observed signals is 2. The mixing matrix is generated using rand function

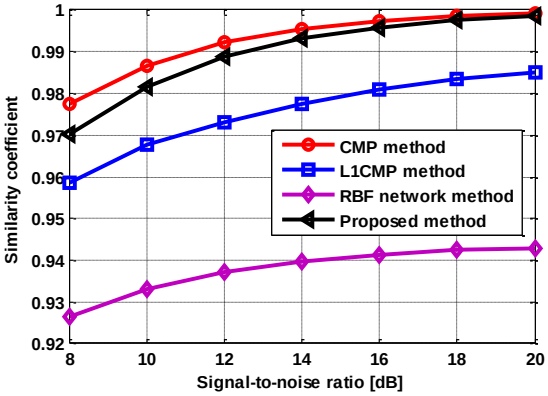
$$\boldsymbol{A}=\begin{bmatrix} 0.3942 & 0.0162 & 0.9100 & 0.2152 & -0.9458 \\ -0.9190 & 0.9999 & -0.4146 & 0.9766 & -0.3246 \end{bmatrix}.$$

When $m=2, n=5$, there are two observed signals, the method proposed in this paper is actually the original shortest path method [7]. The source signals are recovered adopting the method mentioned in this paper and CMP method [15], L1CMP method [15] and RBF network method [16], and the results are shown in Figs. 2 (a)-(c).

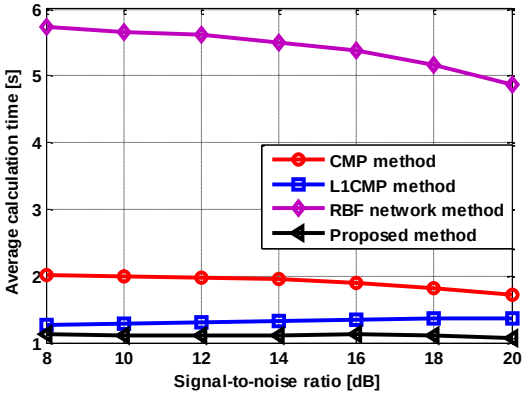
Figure 2 shows that with the increase of signal-to-noise ratio of observed signals, the signal to interference ratio and similarity coefficient of the signals obtained from recovery using different signal recovery algorithms tend to increase, meaning that the signal separation result is acceptable The separation effect adopting the method presented in this paper is slightly lower than CMP method but better than L1CMP and RBF network methods. As for calculation efficiency, it is reflected in Fig. 2 (c) that the calculation time of the method presented in this paper is obviously less than other methods.



(a)



(b)



(c)

Fig. 2. Comparison of the performance and calculation time of different methods when m=2 and n=5. (a) The average signal to interference ratio of separated signals. (b) The average similarity coefficient of separated signals. (c) The average calculation time.

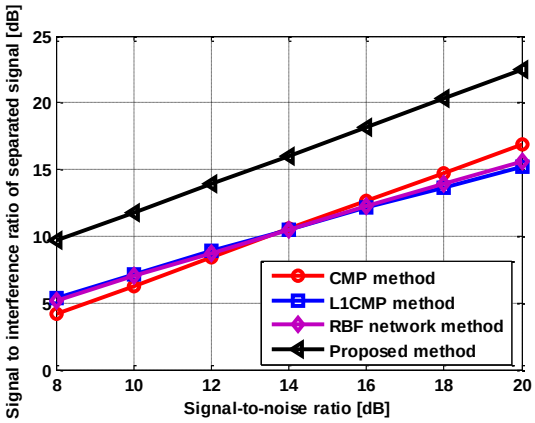
Experiment 2:

The parameters and number of source signals are the same as Experiment 1. The sampling frequency of the receiver is 50MHz and the number of signal sampling points is 10,000. The dimensional of observed signals is three and the mixing matrix is generated using rand function,

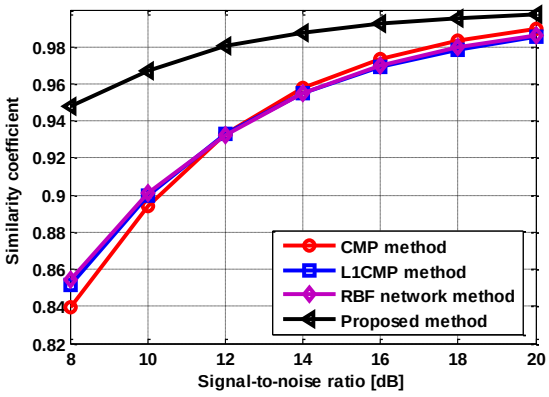
$$\boldsymbol{A}=\begin{bmatrix} -0.1605 & 0.8133 & 0.8729 & 0.3014 & 0.2312 \\ 0.8612 & -0.4460 & 0.4247 & 0.8280 & 0.9358 \\ -0.4823 & 0.3736 & 0.2401 & -0.4728 & -0.2662 \end{bmatrix}.$$

The recovery results of source signals by different methods are shown in Fig. 3.

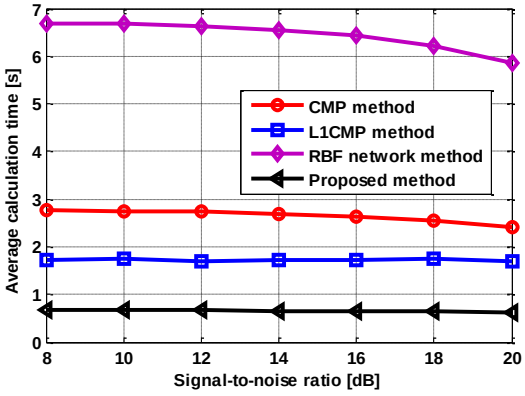
Figure 3 shows that the separation effect adopting the method proposed in this paper is better than the other three methods. The computing efficiency of the method is higher.



(a)



(b)



(c)

Fig. 3. Comparison of the performance and calculation time of different methods when m=3 and n=5. (a) The average signal to interference ratio of separated signals. (b) The average similarity coefficient of separated signals. (c) The average calculation time.

Experiment 3:

There are 7 source signals. Assuming the source signals are sufficiently sparse in time domain, and the

signal types and parameter settings are as follows:

s_1 is a nonlinear FM signal, carrier frequency $f_{c1}=10MHz$, pulse width $t_{r1}=16\mu s$, pulse repetition period $T_{r1}=200\mu s$, instantaneous bandwidth $B_1=10MHz$, pulse start time $t_{01}=0$;

s_2 is a conventional pulse signal, carrier frequency $f_{c2}=8MHz$, pulse width $t_{r2}=15\mu s$, pulse repetition period $T_{r2}=180\mu s$, pulse start time $t_{02}=20\mu s$;

s_3 is a linear FM signal, carrier frequency $f_{c3}=5MHz$, pulse width $t_{r3}=15\mu s$, pulse repetition period $T_{r3}=180\mu s$, pulse start time $t_{03}=40\mu s$, intrapulse bandwidth $B_3=20MHz$;

s_4 is a linear FM signal, carrier frequency $f_{c4}=5MHz$, pulse width $t_{r4}=20\mu s$, pulse repetition period $T_{r4}=180\mu s$, pulse start time $t_{04}=60\mu s$, intrapulse bandwidth $B_4=15MHz$;

s_5 is a sinusoidal phase-modulated signal, carrier frequency $f_{c5}=5MHz$, pulse width $t_{r5}=20\mu s$, pulse repetition period $T_{r5}=200\mu s$, pulse start time $t_{05}=80\mu s$, modulation frequency $f_{a5}=200kHz$, modulation index $a_5=5$;

s_6 is a sinusoidal phase-modulated signal, carrier frequency $f_{c6}=5MHz$, pulse width $t_{r6}=15\mu s$, pulse repetition period $T_{r6}=200\mu s$, pulse start time $t_{06}=100\mu s$, modulation frequency $f_{a6}=200kHz$, modulation index $a_6=2$;

s_7 is a nonlinear FM signal, carrier frequency $f_{c7}=15MHz$, pulse width $t_{r7}=20\mu s$, pulse repetition period $T_{r7}=200\mu s$, intrapulse bandwidth $B_7=5MHz$, pulse start time $t_{07}=115\mu s$.

The sampling frequency of the receiver is 50MHz; the number of sampling points is 10,000 and the dimensional of observed signals is 4. The mixing matrix is generated using rand function,

$$\boldsymbol{A}=\begin{bmatrix} -0.5224 & -0.4859 & 0.1343 & -0.3511 & 0.4795 & 0.0934 & 0.2742 \\ 0.2835 & -0.5464 & -0.4342 & -0.1150 & -0.7562 & -0.6842 & -0.0333 \\ 0.5654 & 0.0297 & 0.7147 & 0.6674 & 0.4023 & 0.6325 & -0.6892 \\ -0.5719 & -0.6815 & 0.5316 & 0.6466 & 0.1906 & 0.3508 & 0.6698 \end{bmatrix}.$$

The recovery results of source signals by different methods are shown in Fig. 4.

Figure 4 shows that the separation effect adopting our method is slightly lower than the other three methods, but the signal to interference ratio of separation and similarity coefficient obtained using our algorithm are sufficient to accurately separate 7 source signals from 4 observed signals, with the law of the computing efficiency is generally the same as Experiment 2.

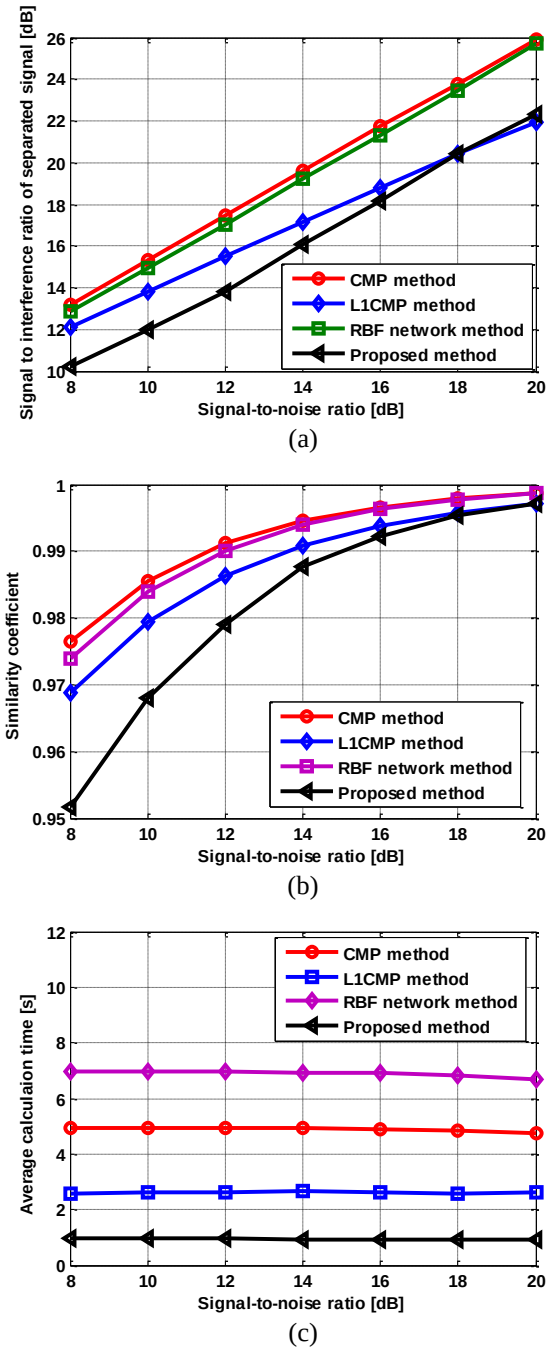


Fig. 4. Comparison of the performance and calculation time of different methods when m=4 and n=7. (a) The average signal to interference ratio of separated signals. (b) The average similarity coefficient of separated signals. (c) The average calculation time.

Experiment 4:

There are 5 source signals, and the source signals are not sufficiently sparse in time domain. Wavelet

transform is used for sparse representation of source signals in this paper. The source signal types and parameter settings are as follows:

s_1 is a traditional pulse signal, carrier frequency $f_{c1}=5MHz$, pulse width $t_{r1}=100\mu s$, pulse repetition period $T_{r1}=400\mu s$, pulse start time $t_{01}=0$;

s_2 is a traditional pulse signal, carrier frequency $f_{c2}=20MHz$, pulse width $t_{r2}=100\mu s$, pulse repetition period $T_{r2}=400\mu s$, pulse start time $t_{02}=0$;

s_3 is a linear FM signal, carrier frequency $f_{c3}=20MHz$, pulse width $t_{r3}=50\mu s$, pulse repetition period $T_{r3}=200\mu s$, pulse start time $t_{03}=100\mu s$, intrapulse bandwidth $B_3=5MHz$;

s_4 is a linear FM signal, carrier frequency $f_{c4}=60MHz$, pulse width $t_{r4}=50\mu s$, pulse repetition period $T_{r4}=200\mu s$, pulse start time $t_{04}=150\mu s$, intrapulse bandwidth $B_4=5MHz$;

s_5 is a sinusoidal phase-modulated signal, carrier frequency $f_{c5}=15MHz$, pulse width $t_{r5}=50\mu s$, pulse repetition period $T_{r5}=200\mu s$, pulse start time $t_{05}=150\mu s$, modulation frequency $f_{a5}=200kHz$, modulation index $a_5=1$.

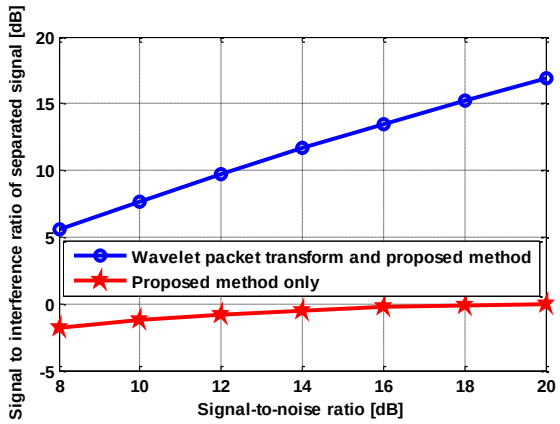
The sampling frequency of the receiver is 200MHz; the number of sampling points is 40,000. The values of two source signals at some moments are not zero, at some moments, only the value of one source signal is non-zero, while the values of other source signals are all zero, that is, the source signals are not sufficiently sparse in time domain. The mixing matrix is generated using rand function,

$$\boldsymbol{A}=\begin{bmatrix} 0.4638 & -0.6693 & -0.9474 & 0.3031 & -0.6117 \\ -0.5711 & -0.4033 & 0.0377 & -0.0828 & -0.7777 \\ 0.6773 & 0.6240 & -0.3177 & -0.9494 & 0.1454 \end{bmatrix}.$$

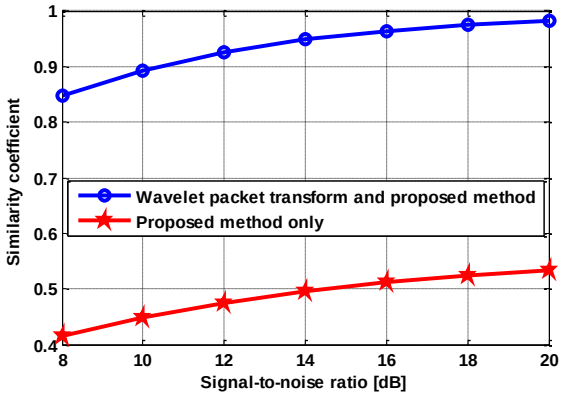
Figure 5 (a) and Fig. 5 (b) show the results by two processing methods, which are adopting the proposed shortest path method based on wavelet transformation and adopting the proposed shortest path method directly. The wavelet basis function is “dmey” and the wavelet is decomposed into 6 layers.

We can see that from Fig. 5, when the source signals are not sufficiently sparse in time domain, time domain separation signals can be obtained by seeking for sparse representation of the source signals using wavelet packet transformation. When compared with the source signal recovery result adopting our method directly in time domain, the recovery effect of the former processing method is obviously better.

The algorithm proposed can be used when there are two or more observed signals. In the shortest path method presented, ideal source signal recovery can be realized with high computing efficiency when the source signal is sufficiently or not sufficiently sparse in time domain. When the source signal is not sufficiently sparse, sparse representation can be used for the observed signals. When the source signal is sufficiently sparse in the transformation domain, the proposed algorithm can be used for source signal recovery.



(a)



(b)

Fig. 5. Performance comparison of different methods when m=3 and n=5. (a) The average signal to interference ratio. (b) The average similarity coefficient.

V. CONCLUSIONS

The shortest path method is a recovery algorithm for underdetermined blind source separation with good effect. It has the advantages of short computational time and high signal recovery accuracy. A new shortest path source signal recovery algorithm is presented based on the defect that the traditional shortest path method can only be used for the case with two observed signals. Underdetermined blind source signal recover can be realized when there are two or more observed signals by employing the proposed algorithm. The feasibility of the algorithm is validated by simulation experiments, and the proposed algorithm has high computing

accuracy and low time complexity.

REFERENCES

[1] Q. Su, Y. H. Shen, Y. M. Wei, and C. L. Deng, "Underdetermined blind source separation by a novel time-frequency method," *International Journal of Electronics and Communications*, vol. 77, pp. 43-49, 2017.

[2] C. C. Wang, Y. H. Zeng, W. H. Fu, and L. D. Wang, "Estimation method for an under-determined mixing matrix based on maximum density point searching," *Journal of Xidian University*, vol. 46, no. 1, pp. 106-111, 2019.

[3] O. Yilmaz, and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1830-1847, 2004.

[4] M. Cobos and J. J. Lopez, "Maximum a posteriori binary mask estimation for underdetermined source separation using smoothed posteriors," *IEEE Trans. Audio, Speech and Language Process.*, vol. 20, no. 7, pp. 2059-2064, 2012.

[5] S. Araki, H. Sawada, R. Mukai, et al., "Underdetermined blind sparse source separation for arbitrarily arranged multiple sensors," *Signal Process.*, vol. 87, no.8, pp. 1833-1847, 2007.

[6] D. Z. Peng and Y. Xiang, "Underdetermined blind source separation based on relaxed sparsity condition of sources," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 809-814, 2009.

[7] P. Bofill, and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Process.*, vol. 81, pp. 2353-2362, 2001.

[8] P. Georgiev, F. Theis, and A. Cichocki, "Sparse component analysis and blind source separation of underdetermined mixtures," *IEEE Trans. Neural Networks*, vol. 16, no. 4, pp. 992-996, 2005.

[9] M. Xiao, S. L. Xie, and Y. L. Fu, "A statistically sparse decomposition principle for under-determined blind source separation," *International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 165-168, Dec. 13-16, 2005.

[10] M. Zhao, "Research on the Theory and Key Problems of Blind Source Separation," *Ph.D. dissertation*, Electronic and Information Dep., South China University of Technology, GuangZhou, China, 2010.

[11] W. H. Fu, J. H. Chen, and B. Yang, "Source recovery of underdetermined blind source separation based on SCMP algorithm," *IET Signal Process.*, vol. 11, no. 7, pp. 877-883, 2017.

[12] W. H. Fu, J. Wei, N. A. Liu, and J. H. Chen, "Algorithm for source recovery in under-determined blind source separation based on plane Pursuit," *Journal of Systems Engineering and Electronics*, vol. 29, no. 2, pp. 223-228, 2018.

[13] W. H. Fu, B. Nong, X. B. Zhou, J. Liu, and C. L. Li, "Source recovery in underdetermined blind source separation based on artificial neural network," *China Communications*, no. 1, pp. 140-145, 2018.

[14] H. Zayyani, M. Babaie-zadeh, F. Haddadi, and C. Jutten, "On the cramér-rao bound for estimating the mixing matrix in noisy sparse component analysis," *IEEE Signal Process. Letters*, vol. 15, pp. 609-612, 2008.

[15] J. H. Chen, "Research on Source Signal Recovery of Underdetermined Blind Source Separation based on Compressed Sensing," *M.S. thesis*, Communication Engineering Dep., Xidian Univ., Xi'an, China, 2015.

[16] W. H. Fu, B. Nong, J. H. Chen, and N. A. Liu, "Source recovery in underdetermined blind source separation based on RBF network," *Journal of Beijing University of Posts and Telecom.*, vol. 40, no. 1, pp. 94-98, 2017.

Chuanchuan Wang received the B.A., M.Sc. and Ph.D. Degrees from Shi Jia Zhuang, Mechanical Engineering College, China, in 2007, 2009, and 2013, respectively.

Since 2014, he has been a Research Assistant with the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System. He is the author of more than 40 articles, and co-authored 21 inventions. His research interests include blind signal processing and effectiveness evaluation theory.

Rui Jia was born in Mengzhou, Henan, China in 1986. He received the B.S. from Zhengzhou University, in 2008 and the Ph.D. degree from Army Engineering University, in 2014.

Since 2015, he has been a Research Assistant with the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System. He is the author of more than 20 articles, and more than 6 inventions. His research interests include blind source separation theory, Electromagnetic compatibility, and reverberation chamber.

Position Deviation Evaluation for UAV Inspecting Overhead Transmission Line Based on Measured Electric Field

Dongping Xiao, Qi Zheng, Jie Lei, and Sheng Liu

State Key Laboratory of Power Transmission Equipment & System Security and New Technology
Chongqing University, Chongqing, 400044, China
xiaodongping@cqu.edu.cn, 20161113031t@cqu.edu.cn, 201811021001@cqu.edu.cn, 20153394@cqu.edu.cn

Abstract ─ Unmanned aerial vehicle (UAV) used in overhead transmission lines (OTLs) inspection are required to fly along a preset path to conduct special tasks. However, UAVs sometimes deviate from the preset path due to positioning error or control error in practice, thereby reducing the quality of the inspection task or even leading to serious line collision accidents. In this study, a method is proposed to evaluate the UAV position deviation from the preset path in real time on the basis of a measurement and analysis on the electric field generated by transmission lines. A new idea is presented to solve the obstacle avoidance problem for UAV in transmission line inspection. To improve evaluation accuracy, the influences of transmission tower and UAV body are considered in the theoretical calculation model for electric field, and the electric field data are preprocessed to diminish the influence of environmental noise and measurement inherent error. To improve the real-time performance of the evaluation algorithm, the dynamic programming and hidden Markov model (HMM) are combined to form a dynamic-hidden Markov model algorithm, in which the parameters of the HMM are determined by the expected maximization parameter estimation and corrected in real time. The feasibility and accuracy of the proposed method are verified by several simulation examples and experiments.

Index Terms ─ Dynamic-hidden Markov model, flight deviation, transmission line inspection, unmanned aerial vehicle (UAV).

I. INTRODUCTION

Using unmanned aerial vehicle (UAV) to inspect overhead transmission lines can improve work efficiency, save manpower and material resources, and overcome the constraints of geographical environment compared with traditional manual inspection. Therefore, transmission line inspection with UAV has been promoted in recent years. To achieve the patrolling objectives and avoid collision accidents, UAVs should fly along preset paths as far as possible. If flight deviates, warning should be provided, and the position should be adjusted in time.

Several methods, such as GPS [1], machine vision ranging [2], ultrasonic ranging, radar ranging [3], optical ranging [4], and artificial potential field method [5,6], are suitable for measuring the distance between the UAV and physical object. However, these methods have limited application scenario, and improving their measurement accuracy is expensive. For example, numerous signal base stations must be built to improve the precision of GPS ranging. Therefore, a convenient and effective method is necessary to evaluate if the flight positions of UAVs deviate from the preset paths.

Many studies [7,8] have reported that AC overhead transmission lines inherently generate power frequency electric fields in their surrounding space and that the electric field intensity at a space observation point is directly related to the distance between the field point and line. On this basis, the present study proposes a new idea, that is, to use the power frequency electric field of the AV overhead transmission lines as the data foundation to evaluate whether and how much the current flight position of a UAV deviates from the preset path. As the electric fields themselves are generated by transmission lines, the electric field measuring device is a single-sided structure, unlike the laser and ultrasonic ranging devices, which requires the producing and transmitting parts of the source signal and the receiving part of the reflection signal. Moreover, power frequency electric field propagates a long distance that the sensor can measure appropriately. The electric field strength is not affected by the ambient light intensity and environment temperature, which avoids the lack of visual and infrared sensing.

The primary idea of the method proposed in this study is to regard the real-time measured electric field at the UAV's positions and the theoretically calculated electric field values at the corresponding positions on the preset path as the basic data. After preprocessing the basic data, the dynamic-hidden Markov probability model (D-HMM) is introduced to evaluate the deviation and drift direction between the current position and preset path of the UAV.

To improve the evaluation accuracy, influences of the transmission tower and UAV body on the electric field are considered in the calculation model. Moreover,

Submitted On: September 26, 2019
Accepted On: April 12, 2020

the theoretically calculated and measured electric field data are preprocessed to eliminate the influence of environmental noise and inherent measurement error before taking them into D-HMM that is used to improve the real-time performance of the evaluation algorithm.

II. IMPROVED CALCULATION MODEL OF ELECTRIC FIELD AND DATA PREPROCESSING

The accuracy of evaluating the flight position deviation of UAV with the proposed method considerably depends on the accurate calculation of the space electric field generated by AC OTLs.

A. 3D electric field calculation model considering actual states of OTL

The charge simulation method (CSM) is recommended to calculate the power frequency electric field generated by AC OTLs in a large space. An improved 3D model is presented in this study to enhance the calculation accuracy.

Figure 1 shows the constituents and layout structure of OTLs with towers. The two ends of a span fixed at Tower A and Tower B can be not at the same height. The origin of coordinate O is set at the projection on the ground of the lowest point of the span. The mathematical model of the catenary conductors can be described as [9,10]:

$$z(x)=\frac{D}{\alpha}\left(\cosh\frac{\alpha x}{D}-\cosh\frac{\alpha}{2}\right)+H(x_{\mathrm{A}}\le x\le x_{\mathrm{B}}), \quad (1)$$

where D is the length of span, α is the horizontal stress coefficient of conductor, and H is the lower height of conductor from the ground. Then,

$$\alpha=\gamma D/\sigma_0, \quad (2)$$

where γ and σ_0 are the specific load and horizontal stress of the conductor, respectively.

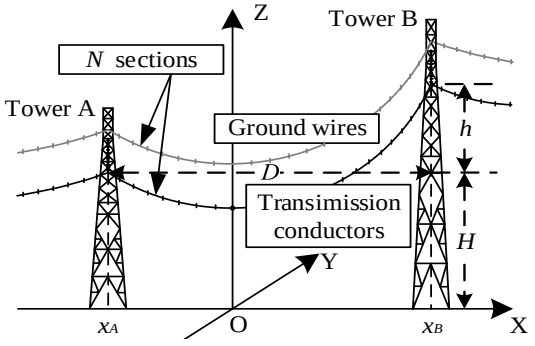


Fig. 1. Constituents and layout structure of the AC OTLs with towers.

If the height of both ends at Tower A and Tower B are equal, then $x_A = -x_B = -D/2$. Meanwhile, if the height difference between two ends is h, then,

$$\begin{cases} x_A=\dfrac{D}{\alpha}\operatorname{arcsinh}\left(\dfrac{\alpha h}{2D}\left(\sinh\dfrac{\alpha}{2}\right)^{-1}\right)-\dfrac{D}{2} \\ x_B=\dfrac{D}{\alpha}\operatorname{arcsinh}\left(\dfrac{\alpha h}{2D}\left(\sinh\dfrac{\alpha}{2}\right)^{-1}\right)+\dfrac{D}{2} \end{cases}. \quad (3)$$

The tower comprises abundant steel bars, which will affect the electric field near them because of electromagnetic induction. A simplified model of the tower is divided into M sections [11]. In the i-th segment, the space coordinates of its head and end are (x_{i1}, y_{i1}, z_{i1}) and (x_{i2}, y_{i2}, z_{i2}), respectively. The coordinates of any point in the middle of this segment is (x_{i0}, y_{i0}, z_{i0}), as shown Fig. 2.

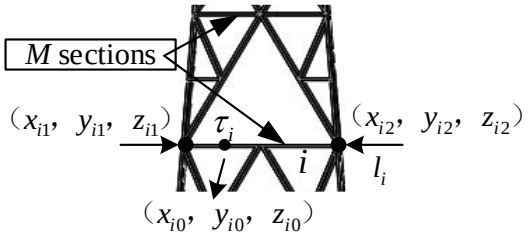


Fig. 2. Model of the tower.

A span of each overhead conductor is divided into N segments, and the coordinate setting for the endpoints of each segment is similar to that shown in Fig. 2. Then, the calculation method is unified for the overhead conductor and tower. Assuming that the length of the i-th analog line charge is l_i and the density of analog line charge is $\dot{\tau}_i$, the potential generated by the analog line charge at the spatial observation point (x, y, z) is denoted by:

$$\dot{\varphi}_i=\frac{l_i}{4\pi\varepsilon_0}\int_0^1\dot{\tau}_i\left(\frac{1}{r}-\frac{1}{r'}\right)dt, \quad (4)$$

where ε_0 is the vacuum dielectric constant; r and r' represent the distance from the source point and its mirror point to the field point, respectively, which are denoted by:

$$\begin{cases} r=\sqrt{\Delta x^2+\Delta y^2+\Delta z^2} \\ r'=\sqrt{\Delta x^2+\Delta y^2+\Delta z'^2} \\ \Delta x=(x_{i2}-x_{i1})t+x_{i1}-x \\ \Delta y=(y_{i2}-y_{i1})t+y_{i1}-y \\ \Delta z=(z_{i2}-z_{i1})t+z_{i1}-z \\ \Delta z'=(z_{i1}-z_{i2})t-z_{i1}-z \end{cases}, \quad (5)$$

where t is the normalized parameter of the distance between points (x_{i0}, y_{i0}, z_{i0}) and (x_{i1}, y_{i1}, z_{i1}), which can be shown as:

$$t=\frac{x_{i0}-x_{i1}}{x_{i2}-x_{i1}}=\frac{y_{i0}-y_{i1}}{y_{i2}-y_{i1}}=\frac{z_{i0}-z_{i1}}{z_{i2}-z_{i1}} \quad (t\in[0,1]). \quad (6)$$

The simulated charge density can be calculated after setting the known conductor potential $\varphi_{il} = U$ and tower potential $\varphi_{it} = 0$ as boundary conditions. Then, the 3D electric field components at the spatial observation point can be obtained:

$$\dot{E}_{ix} = \frac{\partial \dot{\varphi}_i}{\partial x}, E_{iy} = \frac{\partial \dot{\varphi}_i}{\partial y}, E_{iz} = \frac{\partial \dot{\varphi}_i}{\partial z}. \quad (7)$$

Furthermore, numerical calculations are applied to integral Equation (4) and differential Equation (7). Then, the 3D electric field components are given by:

$$\dot{E}_{ix} = \frac{\dot{\tau}_i}{2\pi\varepsilon_0}\left[\frac{a\sqrt{C}+a'\sqrt{A}}{\sqrt{C}\left(2\sqrt{AC}+B\right)} - \frac{a\sqrt{T}+(a+a')\sqrt{A}}{\sqrt{T}\left(2\sqrt{AT}+B+2A\right)}\right]$$

$$\dot{E}_{iy} = \frac{\dot{\tau}_i}{2\pi\varepsilon_0}\left[\frac{b\sqrt{C}+b'\sqrt{A}}{\sqrt{C}\left(2\sqrt{AC}+B\right)} - \frac{b\sqrt{T}+(b+b')\sqrt{A}}{\sqrt{T}\left(2\sqrt{AT}+B+2A\right)}\right], \quad (8)$$

$$\dot{E}_{iz} = \frac{\dot{\tau}_i}{2\pi\varepsilon_0}\left[\frac{c\sqrt{C}+c'\sqrt{A}}{\sqrt{C}\left(2\sqrt{AC}+B\right)} - \frac{c\sqrt{T}+(c+c')\sqrt{A}}{\sqrt{T}\left(2\sqrt{AT}+B+2A\right)}\right]$$

where,

$$\begin{cases} a = x_{i2} - x_{i1} \\ b = y_{i2} - y_{i1} \\ c = z_{i2} - z_{i1} \end{cases}, \begin{cases} d = x_{i1} - x \\ e = y_{i1} - y \\ f = z_{i1} - z \end{cases}, \begin{cases} A = l_i^2 = a^2 + b^2 + c^2 \\ B = 2(ad + be + cf) \\ C = d^2 + e^2 + f^2 \\ T = A + B + C \end{cases}. \quad (9)$$

Finally, the 3D electric field components at the spatial observation point are superimposed the effect of all the simulated charges:

$$\dot{E}_x = \sum \dot{E}_{ix}, \dot{E}_y = \sum \dot{E}_{iy}, \dot{E}_z = \sum \dot{E}_{iz}. \quad (10)$$

The electric field strength can be obtained

$$E = \sqrt{\dot{E}_x^{\,2} + \dot{E}_y^{\,2} + \dot{E}_z^{\,2}}. \quad (11)$$

B. Effect of UAV on the space electric field

The shell material of the UAV used in transmission line patrol and inspection is a composite material with carbon-fiber reinforced resin matrix, which has high electrical conductivity [12,13]. Therefore, the effect of the UAV's body on the space electric field cannot be ignored.

Because the UAV body's shape is complicated and the surface potential is unknown, the electric field analysis cannot be performed with the CSM. The finite element method is suitable for solving partial differential equations on complex regions. Therefore, it is used here.

The simulation model of the UAV is built in the power frequency electric field of transmission lines, and the free tetrahedron element is used to divide the model [14,15]. The boundary conditions of the entire model are initially set. Then, the boundary potential distribution and electric field of each grid element of the UAV model can be calculated with Maxwell equations. In addition, the influence of UAV on electric field distribution can be observed by comparing the electric field distribution in the calculation domain with and without the UAV. The installation position of the electric field measurement sensor on the UAV can be determined as well. In addition, the regularity of UAV's influence also can be used to correct the measured electric field data.

C. Preprocessing of electric field data

To eliminate the influence on electric field measurement from environmental noise and the inherent error of the measuring device, the proposed evaluation method should be made suitable for OTLs with different voltage levels and arrangements. A parameter, namely, electric field ring ratio change rate K (%), is defined to preprocess the electric field data:

$$K_n = \left(E_{n+1} - E_n\right) / \max\{E_n,\ E_{n+1}\} \times 100, \quad (12)$$

where E_n is the electric field at the n-th measurement point, n = 1, 2, 3…. Considering the requirement of the subsequent D-HMM discriminant model, the measured and theoretical electric field data are processed similarly.

III. DEVIATION EVALUATION MODEL

By comparing the measured electric field data with the theoretical electric field data at the same observation points on the preset path, the UAV position deviation from the preset path can be judged to a certain extent. However, the actual physical conditions are difficult to simulate accurately with theoretical models, and random errors exist in the electric field measurement due to the diversity of the geographical environment and meteorological conditions. Therefore, the measured data should be further analyzed and processed, and a practical position matching algorithm with high robustness and good generalization should be studied for the UAV.

A. Background and introduction of D-HMM

The matching model based on probability can effectively overcome these problems and improve evaluation credibility. Therefore, this study uses a non-supervised probability statistical learning model based on time series data, namely, HMM [16,17]. The HMM can be used to determine the implicit parameters of the time series from the observable data sequences. Then, these parameters are used for further analyses, such as pattern recognition and fault diagnosis.

The change process of position and electric field measurement data during UAV patrol and inspection are time-series processes. Furthermore, the current position of the UAV is only related to the previous moment, which satisfies the Markov property [18]. Therefore, the position of UAV can be described by Markov chain.

The hidden here refers to the real position of the UAV, which cannot be obtained directly due to measurement error and must be inferred by the HMM model.

The electric field data are measured in real time during UAV flying. To improve the real-time performance of the model algorithm, this study combines the dynamic programming and HMM to form the D-HMM algorithm. The parameters of the D-HMM are determined with the expectation maximization parameter estimation method and corrected with the dynamic programming method in real time, where the dynamic programming algorithm is used to infer the position sequence of the UAV during flight. Figure 3 shows the main program flow chart of the D-HMM algorithm, in which the observation sequence $\boldsymbol{O}$ and EM parameter estimation and the Viterbi algorithm are three subroutines.

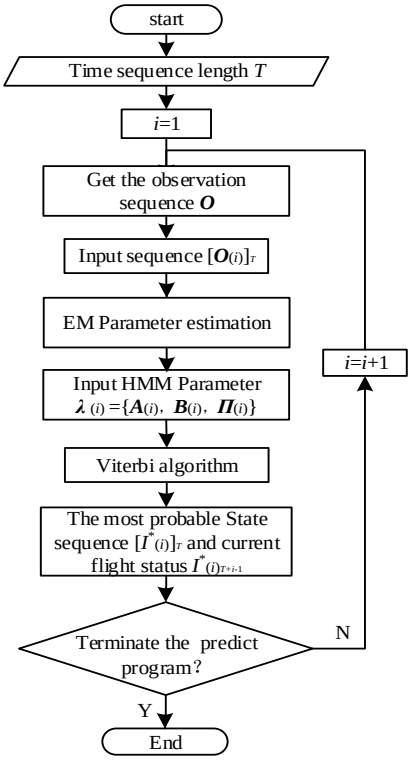


Fig. 3. Flow chart of the D-HMM algorithm.

B. Method of obtaining the observation sequence

Before running the D-HMM algorithm, to simplify the data structure of the algorithm program and improve the running speed of the program, normalizing the measured electric field data with noise is necessary.

The observation sequence $[\boldsymbol{O}_{(i)}]_T$ corresponding to the measured electric field sequence is obtained by comparing the theoretical and measured electric field data after being computed and normalized, where $\boldsymbol{O}_{(i)}$ represents the observation sequence of the i-th iteration.

Initially, the theoretical electric field of the eight-direction offset path around the preset path is shown in Fig. 4. Then, nine theoretical electric field sequences are obtained, where, d = 0,1,2⋯8. They are converted into theoretical K-value sequence $[\boldsymbol{K}_C]_d$ following the method described in Section C of the first part. Furthermore, the time series of measured electric field is also transformed with the same method, and the time series of measured K-value $[\boldsymbol{K}_M]$ is obtained. Then, the initial observation sequence $[\boldsymbol{O}_{(i)}]T$ is obtained by comparing sequence $[\boldsymbol{K}_M]$ and sequence $[K_C]_d$ with the flight velocity v and electric field measurement time interval Δt of the UAV.

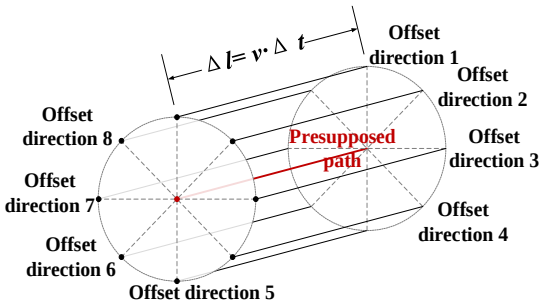


Fig. 4. Schematic of the preset path and offset direction.

C. Model training

After obtaining the input observation sequence, the model should be trained in accordance with the observation sequence to obtain the model parameters.

In practical applications, the data storage of the UAV electric field measurement is limited, the training sample capacity is small, and the data update speed is fast. To realize the rapid training of small sample statistical model, this study uses EM parameter estimation iterative algorithm to obtain HMM parameter set $\boldsymbol{\lambda} = \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\Pi}\}$. $\boldsymbol{\lambda}$ contains three types of parameters, where $\boldsymbol{A} = [a_{ij}]$ is the probability transfer matrix of the real state and a_{ij} represents the probability that the real state changes from i to j. $\boldsymbol{B} = [b_{ik}]$ is the probability matrix of the observed state, and b_{ik} represents the probability of the observed state being k when the real state is i. $\boldsymbol{\Pi} = [\pi_i]$ is the probability vector of the initial state, and π_i represents the probability of the real state being i.

The core steps of the EM algorithm include the two parts of the problem expectation expression and the expectation maximization calculation [19]. In the proposed algorithm model, the expected Q expression is denoted by:

$$\begin{aligned} Q(\boldsymbol{\lambda}_{\boldsymbol{x}}, \boldsymbol{\lambda}_{\boldsymbol{x\text{-}1}}) &= P(\boldsymbol{O}, \boldsymbol{I} \mid \boldsymbol{\lambda}_{\boldsymbol{x\text{-}1}}) \sum_{\boldsymbol{I}} \log \pi_{i(t=1)} \\ &+ P(\boldsymbol{O}, \boldsymbol{I} \mid \boldsymbol{\lambda}_{\boldsymbol{x\text{-}1}}) \sum_{\boldsymbol{I}} \left(\sum_{t=1}^{T-1} \log a_{i(t)\,j(t+1)} \right) \\ &+ P(\boldsymbol{O}, \boldsymbol{I} \mid \boldsymbol{\lambda}_{\boldsymbol{x\text{-}1}}) \sum_{\boldsymbol{I}} \left(\sum_{t=1}^{T} \log b_{i(t)k(t)} \right) \end{aligned} \quad . \qquad (13)$$

The three elements of HMM appear in the three parts of Equation (13). Therefore, only the three parts of the Q function should be maximized to obtain the model parameters:

$$\begin{aligned} \pi_{xi} &= \gamma_{(1)}(i) \\ a_{xij} &= \sum_{t=1}^{T-1}\xi_{(t)}(i,j) / \sum_{t=1}^{T-1}\gamma_{(t)}(i) \\ b_{xik} &= \sum_{t=1,b_t=k}^{T}\gamma_{(t)}(i) / \sum_{t=1}^{T}\gamma_{(t)}(i) \end{aligned}, \tag{14}$$

where,

$$\begin{aligned} \gamma_{(t)}(i) &= P\left(i_{(t)} \mid \boldsymbol{O}, \boldsymbol{\lambda}_{x\text{-}1}\right) \\ \xi_{(t)}(i,j) &= P\left(i_{(t)}, j_{(t+1)} \mid \boldsymbol{O}, \boldsymbol{\lambda}_{x\text{-}1}\right) \end{aligned}. \tag{15}$$

By substituting Equation (15) into Equation (14), the model parameter λ_x calculated by the x-th iteration can be obtained. For one input observation sequence $\boldsymbol{O}$, x iterations are performed. If λ_x converges to λ or reaches the maximum number of iterations, then λ is the output parameter of the model.

D. Dynamic update of HMM parameter

With the movement of the UAV, the track of the UAV will change constantly, and the measured electric field data will be updated constantly. To improve the real-time and accuracy of the algorithm further, the parameters of the algorithm model are continuously corrected with the updating of the input data. In this study, a parameter dynamic update module is added to the HMM algorithm to form the D-HMM algorithm.

The value of the algorithm parameter set $\boldsymbol{\lambda}$ is affected by the observation sequence. Hence, a well-designed dynamic update scheme of the observation sequence can simultaneously realize the dynamic update of the $\boldsymbol{\lambda}$ parameter.

When the algorithm runs for the first time, the observation sequence is recorded as $\boldsymbol{O}_{(1)}$, and the sequence length is T. When the data at time $t = T+1$ is measured, the data at time $t = 1$ is discarded to form a new observation sequence $\boldsymbol{O}_{(2)}$. In this manner, the real-time input sequence group of D-HMM can be obtained. Figure 5 shows the dynamic selection of the input observation sequence.

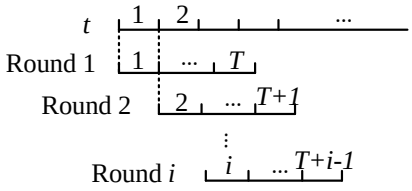


Fig. 5. Dynamic selection of the input observation sequence.

This dynamic selection method of the propulsion input sequence not only preserves the contribution of each measurement data to the calculation of the algorithm parameters during flight but also effectively reduces the sensitivity of the algorithm to the initial parameters. Moreover, this method controls the length of the input sequence, improves the running speed of the algorithm, and enhances the accuracy and timeliness of the discriminant results.

E. Position deviation evaluation

On the basis of the model parameters, the subsequent step is to evaluate whether the UAV deviates from the preset path, particularly the direction and distance of the deviation in a continuous period of time. UAV offset measurement is essentially an optimization analysis problem for a multi-step multi-state process. In this study, the dynamic programming Viterbi algorithm is used to solve the D-HMM model of the known observation sequence $\boldsymbol{O}$ and parameter λ. At this time, the optimal hidden sequence corresponding to the observation sequence is obtained, that is, the true migration of UAV. The core idea of the Viterbi algorithm is that, in the multi-step and multi-choice problem, each step retains the optimal solution of the previous step and finds the optimal path of the transition between multiple steps through the backtracking method [20,21]. Figure 6 shows the algorithm flow.

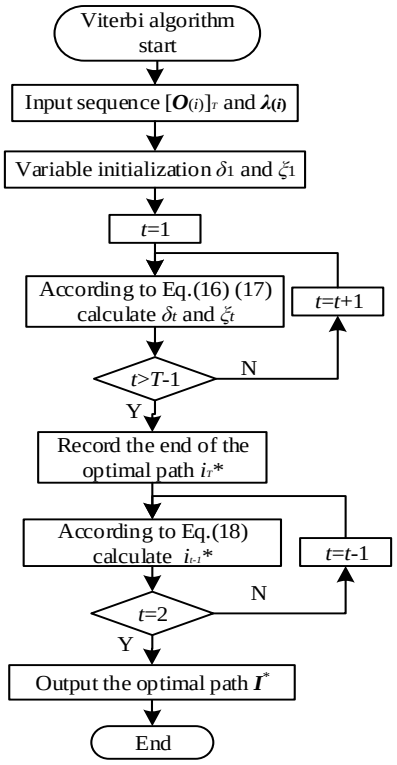


Fig. 6. Flow chart of the Viterbi algorithm.

The variable $\delta_t(i)$ represents the probability of the optimal path in all paths leading to state i at a certain time t. The variable $\Psi_t(i)$ represents the state of the optimal path starting point to state i at time t, the recursive expressions of $\delta_t(i)$ and $\Psi_t(i)$ are as follows:

$$\delta_t(i) = \max_{1\le i,\, j\le 9}[\delta_{t-1}(j)a_{ji}]b_{ik(t)}, \tag{16}$$

$$\psi_t(i) = \arg\max_{1\le i,\, j\le 9}[\delta_{t-1}(j)a_{ji}]. \tag{17}$$

By the forward recursive calculation, the probability P^* corresponding to the last node of the time series and the state I_T^* of the node are obtained. On the basis of the formula,

$$i_t^* = \psi_{t+1}(i_{t+1}^*). \tag{18}$$

The last node is inversely backtracked to obtain the entire optimal path $\boldsymbol{I}^* = [i_1^*, i_2^*, \ldots, i_T^*]$, that is, the UAV deviates from the preset path evaluation result for a period.

IV. EXAMPLE AND EXPERIMENT

To verify the feasibility and validity of the proposed method, a numerical example and experimental scheme are shown here. In the numerical example, assuming that the UAV flight control system is in perfect condition, the real flight path of the UAV coincides with the preset path, but the electric field measurement device has measurement errors. This scheme is designed to verify the effectiveness of the path discrimination algorithm in the case of limited accuracy of the airborne measuring device. The experimental conditions of the experimental verification are set as follows: set up a test site to simulate an experimental platform, set an experimental path that deviates from the preset path, measure the electric field on the path, and input the algorithm model to simulate the discriminant experiment. This scheme is close to the actual situation and aims to verify the feasibility of the algorithm.

A. Numerical verification

The preset path and offset range are set as shown in Fig. 7, and the preset path is numbered as state 0. The four offset states shown in the figure correspond to state numbers 1–4. Figure 8 shows the corresponding K values of the preset paths and offset states.

If the flight plan of UAV is to fly straight from one base tower to the next at a uniform speed, then the flight path is at a horizontal distance of 5 m from the side phase, the height is the same as the ground line suspension, and the total length of the path is 300 m. The electric field is measured every meter of flight. When the UAV reaches the end of the preset path, 301 electric field measurement data are obtained. Then, the corresponding 300 K-values can be calculated. Ideally, the flight state observation sequence of the UAV on the preset path is $O_E = \underbrace{[0,0,0,\ldots\ldots,0,0,0]}_{300*0}$.

The length of the sequence selected by one calculation is set to T=10. To show the UAV flight without offset, the vector representing the probability of actual flight state should be set to $\boldsymbol{\Pi}$ = [1, 0, 0, 0, 0]. The length of the sequence selected by one calculation is set to T=10.

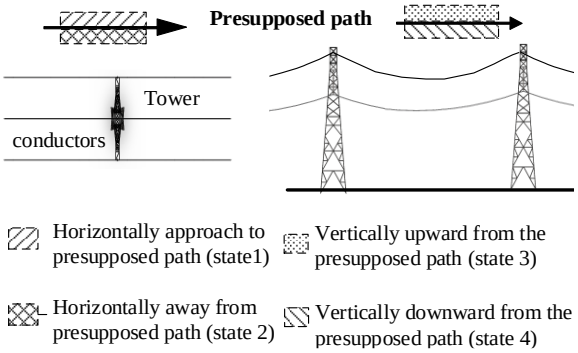


Fig. 7. Schematic of the preset path and offset range (top view and left view).

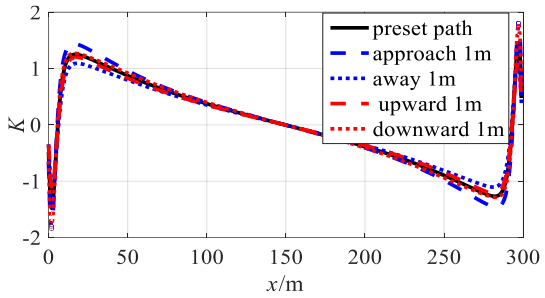


Fig. 8. Preset path and offset value corresponding to the K value.

To show the UAV flight without offset, the vector representing the probability of actual flight state should be set to $\boldsymbol{\Pi}$ = [1, 0, 0, 0, 0]. The random errors of 5%, 10%, 15%, and 20% are superimposed on the electric field value of the preset path to calculate the corresponding K-value. As shown in Fig. 9, the state closest to the K-value is found and forms the observation sequence $\boldsymbol{O}$. The path state I^* is determined by the D-HMM algorithm. Taking the first 10 m of the flight as an example, in one calculation, the measurement error is set to 5%. The input observation sequence is $\boldsymbol{O}_{5\%}$ = [1, 2, 0, 3, 4, 1, 2, 1, 2, 1], and the output path state sequence is $I^*_{5\%}$ = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0]. Therefore, the flight state of UAV is along the preset path without offset in the course of 10 m uniform straight flight from one base tower to the next.

For each of these error ranges, discriminant calculations are performed 30 times using the algorithm model without adding a dynamic parameter update module, and 10 discriminant points are calculated each time. The discriminant error rate of the proposed algorithm is recorded in each error range. Then, the algorithm model of the dynamic parameter update module is used to discriminate the same point. Table 1 shows the results.

Table 1: Accuracy of different error range algorithms

Error Range	Discriminant Point	No Dynamic Parameter Updating Module		Dynamic Parameter Updating Module	
		Number of Errors	Error Rate	Number of Errors	Error Rate
5%	300	25	8.33%	7	2.33%
10%	300	28	9.33%	8	2.67%
15%	300	30	10%	9	3%
20%	300	33	11%	11	3.67%

The D-HMM algorithm with dynamic parameter update module effectively reduces the error rate and improves the effectiveness of the UAV offset state discrimination algorithm when a certain error exists in the electric field measurement.

B. Experimental verification

In the laboratory environment, an experimental platform is simulated. Three copper rods with length of 3 m and diameter of 1 cm are used to simulate the single-return three-phase transmission lines. Wooden supports with height of 2 m are used at both ends of each conductor. The three-phase conductors are arranged horizontally, and the phase spacing is set at 1.5 m. Three wires are connected to 1 kV three-phase transformer through lead wire. The three wires from left to right are A, B, and C phases, respectively. Figure 9 presents the schematic of the simulated experimental platform, and Fig. 10 displays the real object.

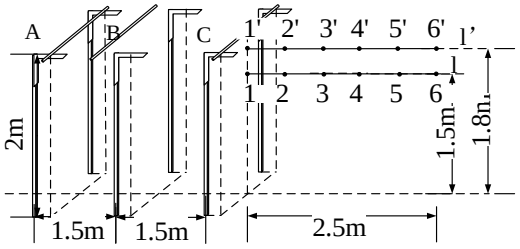


Fig. 9. Schematic of the simulated experimental platform.

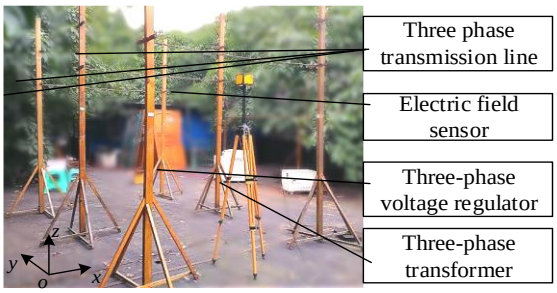


Fig. 10. Physical chart of the simulated experimental platform.

Starting from the midpoint of the side phase conductor, the straight-line *l* in Fig. 9 is set as the preset path of the UAV with a height of 1.5 m. The flight direction is 6→1. An offset straight-line *l'* is set directly above the straight-line *l*, and the height is 1.8 m. The actual flight path of UAV is set to be 6'→5'→4→3→2'→1. Table 2 shows the measured electric field, calculated K value, and discriminant results. Except for the first measurement data without output and the last point has low confidence in discrimination, the remaining points can correctly determine the path offset state. The validity of the proposed algorithm in the laboratory environment is verified.

Table 2: Experimental data and results

Sequence	Electric Field (V/m)	Value of *K*	Discriminant Result
1	43.9		
2	63.7	31.1	5
3	111.6	42.9	4
4	195.2	42.8	3
5	377.5	48.3	2
6	607.2	37.8	1

V. CONCLUSION

This study proposes a method to evaluate the position deviation of UAV from the preset path in real time by using electric field information. Initially, considering the influences of transmission tower and UAV body, a calculation method of spatial electric field distribution is introduced. Then, dynamic programming and HMM are combined to form the D-HMM algorithm for evaluating the deviation of UAV from the preset path. This algorithm can dynamically update model parameters and meet the real-time and accuracy requirements of the deviation evaluation when the UAV is flying with high speed. Finally, the feasibility and effectiveness of the proposed method are verified through numerical calculation and experiments. The application of electric field information in the flight control of line patrol UAV is discussed in depth. This study provides a new thinking of auxiliary UAV flight control and ensures the quality of the inspection task and avoids collision by evaluating the position deviates.

In the future, the work space of multi-functional civil UAVs, such as logistics, photography, and exploration UAVs, and transmission line corridors can have inevitable intersections. The method proposed in this study also provides a new idea for the autonomous obstacle avoidance of generally using UAV.

REFERENCES

[1] J. Kwak and Y. Sung, "Autonomous UAV flight control for GPS-based navigation," *IEEE Access*, vol. 6, pp. 37947-37955, 2018.

[2] E. Zhang and S. Wang, "Plane-space algorithm based on binocular stereo vision with its estimation of range and measurement boundary," *IEEE Access*, vol. 6, pp. 62450-62457, 2018.

[3] P. Peng, L. X. Guo, C. M. Tong, and B. Chen, "A high resolution radar range profile simulator for low flying target above sea surface with multipath effect," *2017 International Applied Computational Electromagnetics Society Symposium (ACES)*, Suzhou, pp. 1-2, 2017.

[4] N. Gageik, P. Benz, and S. Montenegro, "Obstacle detection and collision avoidance for a UAV with complementary low-cost sensors," *IEEE Access*, vol. 3, pp. 599-609, 2015.

[5] Y. Rasekhipour, et al., "A potential field-based model predictive path-planning controller for autonomous road vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1255-1267, May 2017.

[6] G. Wen, et al., "Artificial potential-based adaptive H∞ synchronized tracking control for accommodation vessel," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 7, pp. 5640-5647, July 2017.

[7] W. Yao, et al., "Pioneer design of non-contact synchronized measurement devices using electric and magnetic field sensors," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5622-5630, Nov. 2018.

[8] C. Liao, et al., "Helicopter live-line work on 1000-kV UHV transmission lines," *IEEE Transactions on Power Delivery*, vol. 31, no. 3, pp. 982-989, June 2016.

[9] A. Z. El Dein, M. A. A. Wahab, M. M. Hamada, and T. H. Emmary, "The effects of the span configurations and conductor Sag on the electric-field distribution under overhead transmission lines," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2891-2902, Oct. 2010.

[10] D.-P. Xiao, W. He, F. Yang, Z.-L. Zhang, and J. Tang, "Power frequency electric field calculation and span selection for UHV transmission lines with different meteorological conditions," *High Voltage Eng.*, vol. 35, no. 9, pp. 2081-2086, 2009.

[11] A. Z. El Dein, "Calculation of the electric field around the tower of the overhead transmission lines," *IEEE Transactions on Power Delivery*, vol. 29, no. 2, pp. 899-907, Apr. 2014.

[12] A. Mehdipour, et al., "Advanced carbon-fiber composite materials for RFID tag antenna applications," *Applied Computational Electromagnetics Society Journal*, vol. 25, no. 3, pp. 218-229, 2010.

[13] E. C. Senis, et al., "Enhancement of the electrical and thermal properties of unidirectional carbon fibre/epoxy laminates through the addition of graphene oxide," *Journal of Materials Science*, vol. 54, no. 12, pp. 8955-8970, June 2019.

[14] A. Takei, et al., "High-accuracy electromagnetic field simulation using numerical human body models," *IEEE Transactions on Magnetics*, vol. 52, no. 3, pp. 1-4, Mar. 2016.

[15] F. Aouabed, et al., "Finite element modelling of electric field and voltage distribution on a silicone insulating surface covered with water droplets," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 2, pp. 413-420, Apr. 2018.

[16] N. Asadi, A. Mirzaei, and E. Haghshenas, "Creating discriminative models for time series classification and clustering by HMM ensembles," *IEEE Transactions on Cybernetics*, vol. 46, no. 12, pp. 2899-2910, Dec. 2016.

[17] H. Li, *Statistical Learning Method*, Beijing, China: Tsinghua Univ. Press, 2012.

[18] M. El Yazid Boudaren, et al., "Phasic triplet Markov chains," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2310-2316, Nov. 2014.

[19] F. Guo, O. Wu, Y. Ding, and B. Huang, "A data-based augmented model identification method for linear errors-in-variables systems based on EM algorithm," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 11, pp. 8657-8665, Nov. 2017.

[20] Q. Huang, et al., "Trimming soft-input soft-output Viterbi algorithms," *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2952-2960, July 2016.

[21] C. Soltanpur and J. R. Cruz, "Near optimal Viterbi algorithm for storage channels with linear regressive noise," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 9, pp. 2518-2524, Sep. 2016.

Dongping Xiao was born in Xuanwei, Yunan Province in 1977. She received the B.Sc. degree in Industrial Automation from Chongqing University, Chongqing, China, in 1999, and the M.Sc. and Ph.D. degrees in Electrical Engineering from Chongqing University, in 2004 and 2009, respectively. From 2012 to 2013, she was a Visiting Scholar with Washington State University, USA. Currently, she is an Associate Professor at the College of Electrical Engineering, Chongqing University, China. Her main fields of interests include calculation and simulation of electromagnetic field, electromagnetic measurement and running state monitoring of power transmission equipment.

Electromagnetic Radiation Driven Phase Transition in Silver Telluride-Iron Oxide and Iron Telluride Nano-Composites

Yang Bao [1,3], Wei Zheng [2], Praveen Gurrala [3], Biao Xu [4], Jiming Song [3], and Yue Wu [2]

[1] College of Electronic and Optical Engineering
Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023, China
brianbao@njupt.edu.cn

[2] Department of Chemical and Biological Engineering
Iowa State University, Ames, IA 50011, USA
zhengwei@iastate.edu; yuewu@iastate.edu

[3] Department of Electrical and Computer Engineering
Iowa State University, Ames, IA 50011, USA
praveeng@iastate.edu; jisong@iastate.edu

[4] School of Chemical Engineering
Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China
xubiao@njust.edu.cn

Abstract ─ Temperature-dependent switchability of electrical property has drawn a lot of attention recently. In this article, electromagnetic (EM) radiation is used to induce temperature change in silver telluride and iron telluride nanostructures, which can be potentially used to trigger electrical property changes in them. Due to the low EM absorptivity of silver telluride nanowire, heating it via EM fields is very challenging. To enhance its EM absorptivity, various percentages (by mass) of iron oxide powder are mixed into it. It is verified that, similar to pure silver telluride, the mixture also exhibits a rapid change in electrical conductivity within a certain temperature range. The experiment is designed to characterize the EM absorptivity of three such mixtures with different percentage compositions and an iron telluride nanodisk in the X-band frequency range (8 to 12 GHz). The temperature increase induced by the absorbed EM energy will lead to the rapid change of electrical conductivity by α -type to β -type phase transition in silver telluride and p-type to n-type transition in iron telluride, which can be potentially used to develop EM sensors for applications in communication technology.

Index Terms ─ Electrical switchability, electromagnetic radiation absorption, EM sensor, nano-composites.

I. INTRODUCTION

With the rapid development of wireless communication, the increasing usage of electromagnetic devices results in serious electromagnetic interference (EMI) and electromagnetic compatibility (EMC) problems [1-4]. More seriously, the high density of electromagnetic (EM) radiation caused by wireless communication has been confirmed to have a great effect on the health and safety of humans, such as sleep disturbance, headache, nausea, visual disorders, respiratory problems, and nervous excitation [5].

In consideration of the pollution and threat from EM radiation, many ideas have been proposed to prevent the radiation from being harmful to the humans and communities. EM sensors or switches have since come out. There are some sensors or switches based on amorphous semiconductor [6], superconductor [7], and high-pressure gas arresters [8]. However, semiconductor-based switches have a limited number of operations when protecting circuits against high voltage transients, superconductors require low-temperature cooling, and gas arresters have complicated design with relatively large size [9]. After that, in recent years, significant progress has been achieved in microwave cryogenic electronics operating at liquid nitrogen temperatures [10]. Symmetrical thin-film switches based on current-induced phase transitions from superconductor to normal state [11], and polycrystalline manganite films exhibiting the "electro-resistance" effect [12] were proposed. However, the high speed of operation of these sensors is always accompanied with energy focusing in narrow channels of the film and damage to the protector [10]. To overcome the drawbacks mentioned above, the potential use of silver telluride and iron telluride

Submitted On: November 4, 2019
Accepted On: April 18, 2020

nanostructures in EM sensors are considered, where Joule heating induced by the EM energy absorbed by these materials triggers rapid electrical property changes.

In this article, silver telluride nanowire, which has low EM absorptivity, is blended with iron oxide powder, which has higher ability to absorb EM radiation energy and subsequently convert it into heat or other forms of internal energy [13-17], to improve the mixtures' EM absorptivity while keeping the feature of rapid electrical property changes [18-26]. The EM absorption properties of three mixture samples with different percentages of silver telluride nanowire and iron oxide powder, and a nanodisk made of iron telluride are measured in X band, which is in great demand [27-30]. The good EM absorption properties of these materials can lead to increase in their temperature in environments with high EM radiation. This will eventually cause the rapid change of the electrical conductivity by α -type to β -type phase transition [31-32] in silver telluride and p-type to n-type transition [33-34] in iron telluride, which can be potentially applied to develop EM sensors. The remainder of this paper is organized as follows. Section II describes the experimental details. Section III shows the results and discussion. Finally, the conclusion is drawn in Section IV.

II. EXPERIMENTAL DETAILS

All chemicals are used as received without further purification. Tellurium dioxide (≥99.99%), polyvinylpyrrolidone (PVP, MW~ 40,000), potassium hydroxide, (KOH, ≥90%), hydrazine monohydrate (N_2H_4•H_2O, 78%~82%), and iron oxide (Fe_3O_4, ≥99.99%, 325 mesh) were purchased from Sigma Aldrich, while silver nitrate ($AgNO_3$, ≥99.9%), ethylene glycol (EG, ≥99%) and ethanol (≥95%) were purchased from VWR.

For the synthesis of Ag_2Te_2xAg nanowires, the procedure from our previous publication was followed exactly [32]. Firstly, 9.576 g TeO_2, 12 g PVP, 44.484 g KOH, and 600 ml EG were added to a 1 L glass reactor with magnetic stirring initiated for continuous mixing. The glass reactor was heated to 120°C and 20 ml N_2H_4•H_2O was rapidly injected into the reactor. Here the TeO2 served as a precursor for Tellurium, N_2H_4•H_2O was a reductant to provide electrons with Te, while KOH was added to tune the pH of the solution and the reaction rate. The temperature was maintained at 120°C for 1 hour under nitrogen gas protection. Then, the reaction was stopped and cooled down to room temperature naturally. The tellurium nanowire was washed three times with deionized (DI) water and re-dispersed in 800 ml EG in a 2 L beaker for Ag_2Te synthesis. Alongside, 40.769 g $AgNO_3$ was dissolved in 200 ml EG. Then, the $AgNO_3$/EG solution was added into the 2 L beaker and stirred at room temperature for 2 hours for the conversion from Te into Ag_2Te. In this process, AgNO3 was reduced to metallic Ag and then combined with Te to yield Ag2Te. The as-synthesized Ag_2Te_2xAg nanowires were centrifuged two times with deionized water and washed in an ethanol solution (12.5% of N_2H_4•H_2O) for 24 hours to remove the surfactant. The mixture was then washed twice with ethanol and vacuum-dried.

Ag_2Te-Fe_3O_4 composite sintering: The Ag_2Te nanowires was grounded into loose powder and griddled by a 325-mesh sieve in a nitrogen-filled glovebox. Ag_2Te and Fe_3O_4 powder were mixed evenly by mortar and loaded into graphite die for spark plasma sintering (SPS). The sintering condition was under 40 MPa at 850°C for 5 minutes and then cooled down to room temperature naturally [33].

The EM absorption property was measured using a programmable vector network analyzer (VNA, Agilent E8364) with calibration kit (Agilent 85052D) and cables (GORE Microwave). The samples were placed inside an X band waveguide (Aircom Microwave) to measure the S parameters.

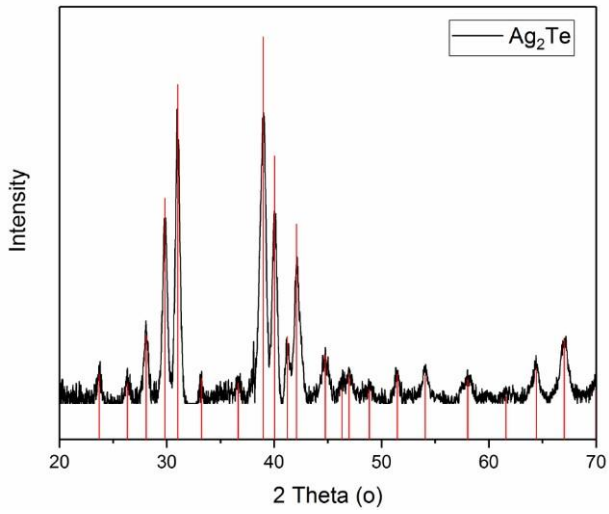
III. RESULTS AND DISCUSSION

The Ag_2Te is analyzed by using X-ray diffraction (XRD) and transmission electron microscopy (TEM). The XRD result, Fig. 1 (a), shows that our synthesized Ag_2Te can be indexed as pure monoclinic phase silver telluride (red lines: JCPDS # 34-1042). Figure 1 (b) is the low-magnification TEM image of Ag_2Te and Fig. 1 (c) is the high-resolution TEM (HRTEM) image of Ag_2Te with its fast Fourier transform (FFT) image. The synthesized Ag_2Te is further proved to be the same monoclinic phase of Ag_2Te (space group is 13).

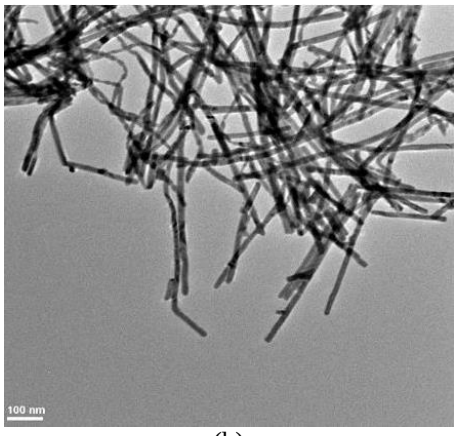
This large-scale solution-synthesis method enables us to sinter several Ag_2Te - Fe_3O_4 composite disks with 1 cm diameter by Spark Plasma Sintering (SPS) and investigate the electrical properties of the resulting mixtures. Figure 2 (a) is the XRD results of Ag_2Te - Fe_3O_4 composites, where all percentages are by weight. There are both silver telluride peaks (JDCPS #34-1042) and iron oxide peaks (JCPDS #19-0692) in all three composite disks. With increasing Ag2Te weight percent, Ag2Te peaks increase while Fe3O4 peaks decrease. The Seebeck coefficient of Ag_2Te - Fe_3O_4 composites has been measured from 40°C to 220°C and is shown in Fig. 2 (b). All these three samples have negative Seebeck coefficient values during the test temperature range which indicate they are n-type semiconductors.

For 25% Ag_2Te-75% Fe_3O_4 and 50% Ag_2Te-50% Fe_3O_4 samples, the absolute Seebeck coefficient increases with temperature. However, the absolute Seebeck coefficient of 75% Ag_2Te -25% Fe_3O_4 decreases first then increases after 125°C. Figure 2 (c) is the electrical conductivity of Ag_2Te - Fe_3O_4 composites. The 25% Ag_2Te -75% Fe_3O_4 sample has the highest electrical conductivity with the value between 140 S/cm to 160 S/cm. The electrical conductivity of all three samples

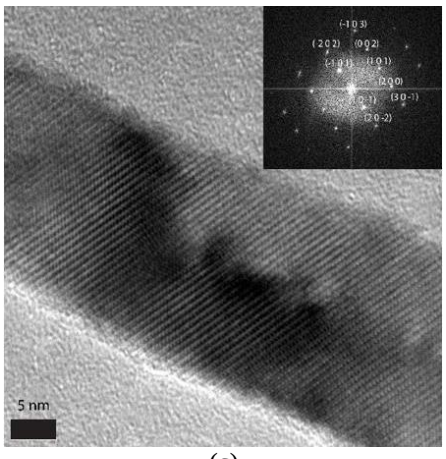
have a dramatic decrease between 140°C to 160°C.



(a)

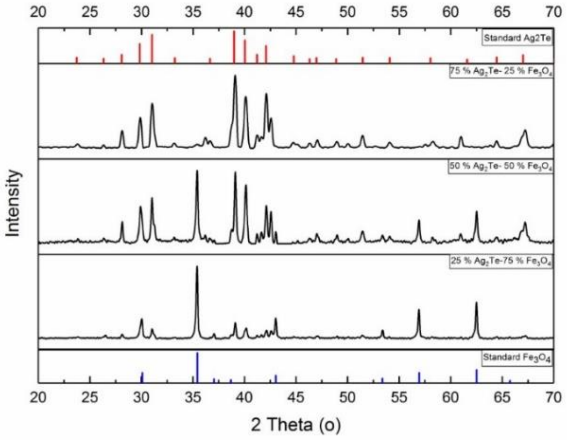


(b)

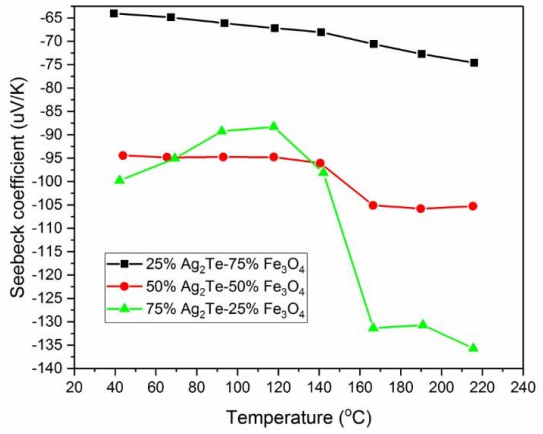


(c)

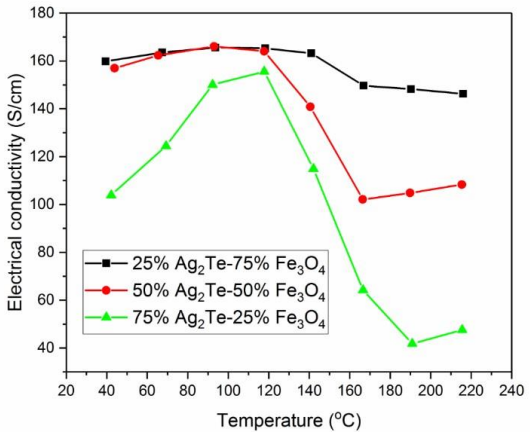
Fig. 1. (a) XRD of Ag_2Te. Red lines: standard silver telluride. (b) TEM image of Ag_2Te. (c) HRTEM of Ag_2Te (Inset: FFT of Ag_2Te).



(a)



(b)



(c)

Fig. 2. (a) XRD results of Ag_2Te - Fe_3O_4 composites. All these percent numbers are weight percentages. (b) Seebeck coefficient of Ag_2Te - Fe_3O_4 composites. (c) Electrical conductivity of Ag_2Te - Fe_3O_4 composites.

As is well known, Ag_2Te changes from low-temperature monoclinic phase (α phase) to high-temperature cubic phase (β phase) at 150°C [35-36].

During phase transition from α phase to β phase, the energy band gap of Ag_2Te increases, which in turn decreases the carrier concentration (n) while the electron mobility (v) fluctuates a little bit. As a result, the electrical conductivity changes from a highly conductive state (α phase) to a poorly conductive state (β phase) based on the equation σ=nev [35, 37, 38]. This phase transition property may explain the dramatic changes in Figure 2b and 2c, especially for the sample of 75% Ag_2Te-25% Fe_3O_4. As to the $FeTe_2$, it shows a temperature-dependent reversible and reproducible switching behavior between p-type to n-type conduction [33].

Magnetite is one of the most important resources for production of iron via microwave heating because of its good coupling with microwaves and three mechanisms including Joule loss, dielectric loss, and magnetic loss would contribute to microwaves heating [39]. Thus, the magnetite is mixed with the silver telluride nanowire to increase the EM absorption of the composites. Figure 3 shows the EM absorption of the three Ag_2Te - Fe_3O_4 composite disks with different mass percentage compositions (75% Fe_3O_4-25% Ag_2Te, 50% Fe_3O_4-50% Ag_2Te, 25% Fe_3O_4-75% Ag_2Te), and the iron telluride nanodisk in X band. The measurement was performed by placing the disks in an X-band waveguide with the disk face parallel to the waveguide cross-section to ensure the same experiment condition. The input power p_0 is 0 dBm. All disks are 1.6 mm thick. The EM absorption is calculated from the measured S parameters and is given by $\left(1-\left|S_{11}\right|^2-\left|S_{21}\right|^2\right)p_0$ [40-42].

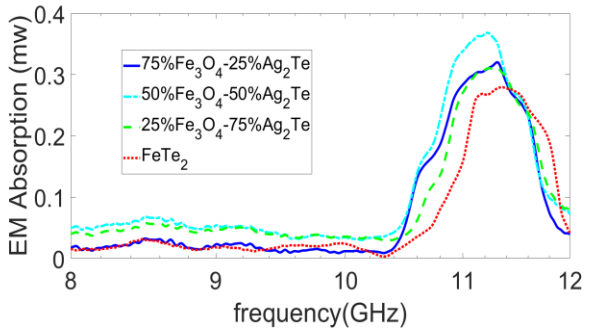


Fig. 3. EM absorption of samples in X band.

The maximum EM absorption power measured is 0.37 mW, which occurs at 11.2 GHz with the composite made of 50% Fe3O4-50% Ag2Te. This composite also shows better EM absorption performance over others in the frequency range from 8 GHz to 11.4 GHz. For the composites made of 25% Fe_3O_4-75% Ag_2Te, 75% Fe_3O_4-25% Ag_2Te, and the iron telluride, the maximum absorption power is 0.31 mW at 11.31 GHz, 0.32 mW at 11.29 GHz and 0.28 mW at 11.35 GHz, respectively. At least 30% power is absorbed by the samples at around 11 GHz. The potential sensors can absorb high-power EM radiation and convert it to heat, which results in a temperature rise given by $\Delta T = Q/C$, where Q is the amount of the heat absorbed and C is the heat capacity [43]. The temperature rise eventually leads to rapid change in electrical conductivity by α-type to β-type phase transition for Ag_2Te - Fe_3O_4 samples as shown in Fig. 2, and p-type to n-type transition in iron telluride with reversible and reproducible switching behavior [33]. A multi-physics simulation package is needed to include electromagnetic modeling of energy absorptivity of the mixtures at different frequencies, thermal modeling of temperature changes, and chemical modeling of electrical conductivity changes from α-type to β-type phase transition in silver telluride and p-type to n-type transition in iron telluride. The electromagnetic radiation driven phase transitions in Ag_2Te - Fe_3O_4 and $FeTe_2$ nano-composites can be potentially applied to develop EM sensors.

IV. CONCLUSIONS

It is shown in this work that the iron telluride and mixture composites made of silver telluride nanowire and iron oxide powder have good EM absorption in X band and exhibit rapid changes of electrical property in a certain temperature range. The composite made of 50% Fe3O4-50% Ag2Te has the maximum EM absorption power at 11.2 GHz which leads to the temperature rise. In turn, the temperature rise causes a rapid change in the electrical conductivity by α-type to β-type phase transition in the composite mixtures and p-type to n-type transition in iron telluride, which can be potentially exploited for developing EM sensors.

V. ACKNOWLEDGEMENTS

This work is supported in part by Office of Naval Research (Award number N00014-16-1-2066), in part by the IU Program of the Center for Nondestructive Evaluation at Iowa State University, and in part by China Scholarship Council.

REFERENCES

[1] G. B Sun, B. X Dong, *et al.*, "Hierarchical dendrite-like magnetic materials of Fe3O4, γ-Fe2O3, and Fe with high performance of microwave absorption," *Chemistry of Materials*, vol. 23, pp. 1587-1593, 2011.

[2] W. M. Zhu, L. Wang, *et al.*, "Electromagnetic and microwave-absorbing properties of magnetic nickel ferrite nanocrystals," *Nanoscale*, vol. 3, pp. 2862-2864, 2011.

[3] Z. Piotrowski, *et al.*, "Electromagnetic compatibility

of the military handset with hidden authorization function based on MIL-STD-461D results," *Progress in Electromagnetics Research*, vol. 4, pp. 566-570, 2008.

[4] G. Y. Slepyan, A. Boag, V. Mordachev, *et al*., "Nanoscale electromagnetic compatibility: Quantum coupling and matching in nanocircuits," *IEEE Transactions on Electromagnetic Compatibility*, vol. 57, no. 6, pp. 1645-1654, Dec. 2015.

[5] D. Sun, Q. Zou, G. Qian, *et al*., "Controlled synthesis of porous Fe3O4-decorated graphene with extraordinary electromagnetic wave absorption properties," *Acta Materialia*, vol. 61, pp. 5829-5834, 2013.

[6] V. Balchiunas, S. Baleivichius, and A. Deksnys, "Size effect of nanosecond switching in submicron coplanar structures based on amorphous semi-conductors," *Lietuvos Fizikos Rinkinys*, vol. 4, no. 405, 1991.

[7] S. Balevičius, *et al*., "Ultrafast electrical super-conducting to normal states switching in Y-Ba-Cu-O and Bi-Sr-Ca-Cu-O microstrips," *MRS Online Proceedings Library Archive*, vol. 275, pp. 1-5, 1992.

[8] T. Kikel, L. Altgilbers, *et al*., "Plasma limiters," in *AIAA Plasmadynamics and Laser Conference*, vol. 98, no. 2564, 1998.

[9] L. Altgilbers, *et al*., "Fast protector against EMP using electrical field induced resistance change in La/sub 0.67/Ca/sub 0.33/MnO/sub 3/thin films," *Pulsed Power Plasma Science*, vol. 2, pp. 1782-1785, 2001.

[10] S. Balevicius, *et al*., "Fast protector against EMP using thin epitaxial and polycrystalline manganite films," *IEEE Electron Device Letters*, vol. 32, pp. 551-553, 2011.

[11] S. Balevicius, *et al*., "EMP effects on high-T/sub c/superconducting devices," *IEEE transactions on Applied Superconductivity*, vol. 14, pp. 112-118, 2004.

[12] N. Žurauskienė, "Thin manganite films for fast fault current limiter applications," *Thin Solid Films*, vol. 515, pp. 576-579, 2006.

[13] Y. Ding, L. Zhang, *et al*., "Electromagnetic wave absorption in reduced graphene oxide functionalized with Fe3O4/Fe nanorings," *Nano Research*, vol. 9, pp. 2018-2025, 2016.

[14] X. Li, X. Yin, C. Song, *et al*., "Self‐assembly core–shell graphene‐bridged hollow MXenes spheres 3D foam with ultrahigh specific EM absorption performance," *Advanced Functional Materials*, vol. 28, no. 41, p. 1803938, Aug. 2018.

[15] M. I. Hossain, M. R. I. Faruque, and M. T. Islam, "A new design of cell phone body for the SAR reduction in the human head," *Applied Computational Electromagnetics Society Journal*, vol. 30, no. 7, pp. 792-798, July 2015.

[16] G. Wang, X. Peng, L. Yu, *et al*., "Enhanced microwave absorption of ZnO coated with Ni nanoparticles produced by atomic layer deposition," *Journal of Materials Chemistry A*, vol. 3, no. 6, pp. 2734-2740, Dec. 2015.

[17] N. O. Parchin, M. Shen, and G. F. Pedersen, "Small-size tapered slot antenna (TSA) design for use in 5G phased array applications," *Applied Computational Electromagnetics Society Journal*, vol. 32, no. 3, pp. 193-202, Mar. 2017.

[18] Y. V. Pershin and D. V. Massimiliano, "Memory effects in complex materials and nanoscale systems," *Advances in Physics*, vol. 60, pp. 145-227, 2011.

[19] A. D. Smith, K. Elgammal, F. Niklaus, *et al*., "Resistive graphene humidity sensors with rapid and direct electrical readout," *Nanoscale*, vol. 7, pp. 19099-19109, Oct. 2015.

[20] J. Yang, D. Joshua, B. Strukov, and R. S. Duncan, "Memristive devices for computing," *Nature Nanotechnology*, vol. 8, no. 13, 2013.

[21] K. Ohta, Y. Kuwayama, K. Hirose, *et al*., "Experimental determination of the electrical resistivity of iron at Earth's core conditions," *Nature*, vol. 534, pp. 95-98, June 2016.

[22] O. Sato, "Dynamic molecular crystals with switchable physical properties," *Nature Chemistry*, vol. 8, pp. 644, 2016.

[23] H. Zhao and J. Bai, "Highly sensitive piezo-resistive graphite nanoplatelet–carbon nanotube hybrids/polydimethylsilicone composites with improved conductive network construction," *ACS Applied Materials & Interfaces*, vol. 7, no. 18, pp. 9652-9659, Apr. 2015.

[24] H. Layssi, P. Ghods, A.R. Alizadeh, *et al*., "Electrical resistivity of concrete," *Concrete International*, vol. 37, no. 5, pp. 41-46, May 2015.

[25] F. Li, *et al*., "Phase-transition-dependent conductivity and thermoelectric property of silver telluride nanowires," *The Journal of Physical Chemistry C*, vol. 112, pp. 16130-16133, 2008.

[26] F. Messerschmitt, M. Kubicek, and J. L. M. Rupp, "How does moisture affect the physical property of memristance for anionic–electronic resistive switching memories?," *Advanced Functional Materials*, vol. 25, no. 32, pp. 5117-5125, July 2015.

[27] C. Guo, X. Shang, J. Li, *et al*., "A lightweight 3-D printed X-band bandpass filter based on spherical dual-mode resonators," *IEEE Microwave and Wireless Components Letters*, vol. 26, no. 8, pp. 568-570, Aug. 2016.

[28] H. Li, G. Wang, H. X. Xu, *et al*., "X-band phase-gradient metasurface for high-gain lens antenna application," *IEEE Transactions on Antennas and*

Propagation, vol. 63, no. 11, pp. 5144-5149, Nov. 2015.

[29] M. Han, *et al*., "Ti3C2 MXenes with modified surface for high-performance electromagnetic absorption and shielding in the X-band," *ACS Applied Materials & Interfaces*, vol. 8, pp. 21011-21019, 2016.

[30] J. Zhang, H. Yang, and H. Liang, "Band-notched split-ring resonators loaded monopole antenna for ultrawideband applications," *Applied Computational Electromagnetics Society Journal*, vol. 28, no. 2, pp. 137-142, Feb. 2013.

[31] A. Qin, *et al*., "Silver telluride nanotubes prepared by the hydrothermal method," *Inorganic Chemistry*, vol. 46, pp. 7403-7409, 2007.

[32] H. Yang, *et al*., "Composition modulation of Ag2Te nanowires for tunable electrical and thermal properties," *Nano Letters*, vol. 14, pp. 5398-5404, 2014.

[33] W. Zheng, *et al*., "Solution-phase synthesized iron telluride nanostructures with controllable thermally triggered p-type to n-type transition," *Nanoscale*, vol. 10, pp. 20664-20670, 2018.

[34] J. H. Lee, J. H. Kim, and S. S. Kim, "CuO–TiO2 p–n core–shell nanowires: Sensing mechanism and p/n sensing-type transition," *Applied Surface Science*, vol. 448, pp. 489-497, Aug. 2018.

[35] S. Aliev, Z. Agaev, and E. Zul'figarov, "Charge transport in silver chalcogenides in the region of phase transition," *Semiconductors*, vol. 41, pp. 1027-1032, 2007.

[36] Y. Pei, N. A. Heinz, and G. J. Snyder, "Alloying to increase the band gap for improving thermoelectric properties of Ag2Te," *Journal of Materials Chemistry*, vol. 21, pp. 18256-18260, 2011.

[37] S. Aliev, "Hysteresis in Ag2Te near and within the phase transition region," *Semiconductors*, vol. 38, pp. 796-799, 2004.

[38] F. Aliev, "Phase transition of Ag-Enriched Ag2Te," *Inorganic Materials*, vol. 38, pp. 995-997, 2002.

[39] A. Amini, *et al*., "Effect of particle size and apparent density on the initial stages of temperature increase during the microwave heating of Fe3O4," *Powder Technology*, vol. 338, pp. 101-109, 2018.

[40] R. S. Rao, *Microwave Engineering*. PHI Learning Pvt. Ltd, 2015.

[41] S. Luo, Y. Li, Y. Xia, *et al*., "A low mutual coupling antenna array with gain enhancement using metamaterial loading and neutralization line structure," *Applied Computational Electromagnetics Society Journal*, vol. 34, no. 3, pp. 411-418, Mar. 2019.

[42] P. Ferrazzoli, L. Guerriero, and D. Solimini, "Numerical model of microwave backscattering and emission from terrain covered with vegetation," *Applied Computational Electromagnetics Society Journal*, vol. 6, no. 1, pp. 175-191, June 1991.

[43] W. J. Parker, *et al*., "Flash method of determining thermal diffusivity, heat capacity, and thermal conductivity," *Journal of Applied Physics*, vol. 32, pp. 1679-1684, 1961.

Adaptive Sparse Array Beamforming Using Correntropy Induced Metric Constrained Normalized LMS Algorithm

Zehua Dai [1,3], Longxiang Guo [1,3,*], Jingwei Yin [1,3], Yingsong Li [2], and Kun Guo [1,3]

[1] Acoustic Science and Technology Laboratory
Harbin Engineering University, Harbin 150001, China
heu503@hrbeu.edu.cn*

[2] College of Information and Communication Engineering
Harbin Engineering University, Harbin 150001, China

[3] Key Laboratory of Marine Information Acquisition and Security, Harbin Engineering University
Ministry of Industry and Information Technology, Harbin 150001, China

Abstract ─In order to further exploit the sparseness of antenna array and speed up the convergence of constrained normalized LMS (CNLMS) algorithm, maintaining good beam pattern performance and better output signal-to-interferences-plus-noise ratio (SINR), a new method with approximation l_0-norm constraint is proposed to improve CNLMS algorithm, and its derivation process is given in detail. In this newly proposed algorithm, the correntropy induced metric (CIM) is used to approximate the l_0-norm, which is considered construct a new cost function to fully exploit the sparsity of the antenna array and reduced the number of active array elements. Using the CIM penalty, the proposed CIM-based CNLMS (CIM-CNLMS) algorithm is derived in detail, where the Lagrange multiplier method is utilized to solve the cost function of the proposed CIM-CNLMS algorithm, and the steepest descent principle is considered to obtain the update equation. The computer simulation results demonstrate that compared with other CLMS algorithms, the new algorithm obtains better performance, which greatly reduces the proportion of active array elements in the thinned antenna array. Simultaneously, the new algorithm has excellent beam pattern performance and better SINR performance with faster convergence speed and more stable mean square error.

Index Terms ─ adaptive array beamforming, CNLMS algorithm, correntropy induced metric, l_0-norm constraint, sparse arrays.

I. INTRODUCTION

Adaptive beamforming, an essential and elementary problem in array signal processing, has attracted great attention in many applications, including sonar, radar, wireless and mobile communications, seismic sensing, biomedical engineering [1-2]. Moreover, the Linear Constrained Minimum Variance (LCMV) algorithm proposed by Frost [3] is a well-known algorithm, which gives a beam in the desired direction and forms a null in the direction of arrival (DOA) of the interfering signal. The LCMV algorithm minimizes the energy output in order to minimize the interference of the array output and the noise signal energy, maintaining a constant gain in the desired direction of the viewing direction. The adaptive beamforming algorithm adjusts the weight vector of the antenna array to match the desired signal as well as the interfering signal as a function of time. The CNLMS algorithm is considered as the normalization form of the LCMV algorithm, in which the array elements can be adjusted in real time [4].

In some applications, such as radar, sonar, in order to achieve the desired performance, we need to use many array elements to achieve the goal. However, if the array has too many array elements, the array requires a lot of operations and huge amounts of energy. Consequently, the application of many existing beamforming algorithms will cost a large amount of energy consumption in the antenna array, and the antenna array needs to provide a strong calculate ability, which make the antenna array increase the cooling equipment with superior performance, resulting in complex antenna equipment and high cost.

Till now, the existing algorithms solve the mentioned problems above. With the development of sparse signal processing technology, it becomes a new exploration direction that the algorithm of sparse signal processing is applied to design beamforming algorithms to make the weight coefficient of the array element toward sparse. In recent years, along with the development of Compressed Sensing [5], many works have been done in the field of sparse signal processing

Submitted On: October 17, 2019
Accepted On: April 5, 2020

[6-15]. In these efforts, LASSO [16] and some new LMS-based algorithms are developed for sparse system identification [6-9]. For example, the zero-attractive least mean square (ZA-LMS) and re-weighted ZA-LMS (RZA-LMS) algorithms have been proposed for this purpose. The ZA-LMS algorithm uses a l_1-norm penalty on the LMS cost function, which gives a zero attractor in the iterations. The RZA-LMS algorithm utilizes a re-weighted zero attractor to further improve the ZA-LMS's performance.

Inspired by LASSO and the sparse LMS algorithms, the l_1-norm constrained LMS (l_1-CLMS) algorithm was proposed in [17]. After that, the new normalized version of l_1-CLMS (l_1-CNLMS) [18] algorithm and its re-weighted version (l_1-WCNLMS) [18] with superior performance have been proposed to make the antenna array element coefficients toward sparse, and the algorithm achieved good convergence performance. Recently, many re-weighted l_1-norm and l_p-norm penalties have been proposed and considered in [19-21]. And the l_0-CNLMS algorithm, which applies the penalty of approximate l_0-norm to adaptive beamforming, has been proposed in [22]. However, its computations are high because of the exponentiation operation.

To better exploit the sparse characteristics of the antenna array, and to fully utilize the advantages of CIM theory for calculating the number of non-zero entries in the array weight vector, a new CNLMS algorithm is proposed by utilizing the CIM theory to reduce the number of active array elements under the framework of adaptive beamforming. In our proposed algorithm, CIM theory is to construct a modified CNLMS cost function that implements a zero attractor in CNLMS's iterations with the help of Gaussian kernel theory, which is named as CIM-CNLMS. As a result, CIM acts as a l_0-norm to help to speed up convergence and reduce MSE of the CNLMS. The simulation results demonstrate that the proposed CIM-CNLMS algorithm can improve the sparseness of the antenna array and reduce the MSE compared with the l_1-WCNLMS algorithm. That is to say, the proposed algorithm has better superiority than the l_1-WCNLMS algorithm in terms of sparsity and MSE characteristics. The algorithms presented in this paper have potential applications in radar, sonar, and 5G antenna arrays.

II. THE ARRARY PROCESSING FUNDAMENTALS

As portrayed in Fig. 1, a planar antenna array model consisting of M omnidirectional antenna elements spaced by half-wavelength is utilized to discuss adaptive beamforming algorithms, where λ represents the wavelength of the electromagnetic wave actually used. Assuming (L+1) narrowband signals are received by the antenna elements, with one signal of interest (SOI) and L interference signals. We assume that the horizontal azimuth of the SOI is θ_{SOI} and the horizontal azimuth of the L interference signals is θ_p (p=1, 2, ..., L). Besides, the zenith of SOI is defined as $\emptyset_{SOI}$ and the zeniths of L interference signals are defined as $\emptyset_p$ (p=1, 2, …, L). The received signal at the n-th snapshot is expressed as:

$$\mathbf{x}(n) = \mathbf{a}_c\mathbf{s}(n) + \sum_{l=1}^{L}\mathbf{a}_l\mathbf{i}_l(n) + \boldsymbol{\eta}(n), \quad (1)$$

where $\mathbf{a}$, $\mathbf{s}(n)$, $\mathbf{i}_l(n)$ and $\boldsymbol{\eta}(n)$ are the steering matrix related to the SOI and interference signal, the complex signal envelope vector as well as zero-mean Gaussian white noise vector, respectively. We assume that the SOI, interfering signals, and noise are statistically independent of each other.

Under these assumptions, the SINR of the beamformer is calculated using the following equation:

$$\mathbf{SINR}_{out} = \frac{\mathbf{w}^{\mathrm{H}}\mathbf{R}_s\mathbf{w}}{\mathbf{w}^{\mathrm{H}}\mathbf{R}_\xi\mathbf{w}}, \quad (2)$$

where $\mathbf{w}$ represents the weight coefficient vector of a M×1 dimensional, and $\mathbf{R}_\xi$ represents the covariance matrix of the interfering signal plus noise, $\mathbf{R}_s$ represents the covariance matrix of the SOI. Then, we have:

$$\begin{cases} \mathbf{R}_s = \mathbf{a}_c\mathbf{s}(n)\mathbf{s}(n)^{\mathrm{H}}{\mathbf{a}_c}^{\mathrm{H}}, \\ \mathbf{R}_\xi = \mathbf{A}_k\mathbf{I}(n)\mathbf{I}(n)^{\mathrm{H}}{\mathbf{A}_k}^{\mathrm{H}} + \boldsymbol{\eta}(n)\boldsymbol{\eta}(n)^{\mathrm{H}}, \end{cases} \quad (3)$$

where $\mathbf{A}_k$ represents the steering matrix of L interfering signals, and $\mathbf{I}$ represents the interfering signal matrix composed of L interfering signals.

The output signal $\mathrm{y}(n)$ of the planar array at the nth snapshot can be expressed as:

$$\mathrm{y}(n) = \mathbf{w}^{\mathrm{H}}\mathbf{x}(n). \quad (4)$$

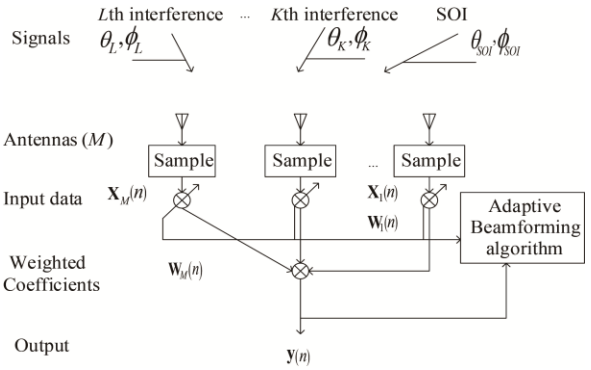


Fig. 1. Beamforming model.

III. THE CLMS ALGORITHM AND THE CNLMS ALGORITHM

A. The CLMS algorithm

The specific results of the LCMV algorithm are presented in [3], in which the weight vector representation of the LCMV algorithm is expressed as:

$$\mathbf{w}_{\mathrm{opt}} = \mathbf{R}^{-1}\mathbf{C}(\mathbf{C}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{C})^{-1}\mathbf{F}, \quad (5)$$

where $\mathbf{R}$ is the covariance matrix of the array input data and is defined as $E(\mathbf{X}\mathbf{X}^{\mathrm{H}})$, $\mathbf{C}$ is the constraint matrix of

the array, and **F** denotes the constraint vector whose elements are related to the signal and interference of interest, $(\cdot)^{\mathrm{H}}$ is the conjugate transpose operator.

The CLMS algorithm adopts an adaptive filtering technique, which can effectively increase the gain of the SOI and at the same time better attenuates the interference signals from other directions.

The following variables are defined as: $e_n \in \mathbb{C}$, $\mathbb{C}$ represents the set of complex numbers, e_n represents the estimation error of the adaptive filter, $\mathbf{w} \in \mathbb{C}^M$, $\mathbf{w}$ represents the vector of the coefficients, $\mathbf{x}_n \in \mathbb{C}^M$, $\mathbf{x}_n$ represents a vector composed of input signals, $d_n \in \mathbb{C}$, d_n represents the desired signal. Using the minimum mean square error criterion, the linear constraint minimum problem can be expressed mathematically as:

$$\min_{\mathbf{w}} E[|e_n{}^2|] \quad \text{s.t. } \mathbf{C}^{\mathrm{H}}\mathbf{w} = \mathbf{F}, \tag{6}$$

where $e_n = d_n - \mathbf{w}^{\mathrm{H}}\mathbf{x}_n$. $\mathbf{C}$ represents an $M \times (L+1)$ constraint matrix, and $\mathbf{F}$ represents a corresponding constraint vector containing (L+1) elements.

To solve this mathematical problem, the Lagrange multiplier is used to construct a cost function, which is given by:

$$\mathrm{G}(n) = E[|e_n|^2] + \boldsymbol{\gamma}^{\mathrm{H}}(\mathbf{C}^{\mathrm{H}}\mathbf{w} - \mathbf{F}), \tag{7}$$

where $\boldsymbol{\gamma}$ represents a Lagrange multiplier. To solve this cost function, the gradient descent principle is utilized for getting the optimal solution of the cost function.

Firstly, the two sides of the equation (7) are derived with respect to $\mathbf{w}$, and the following equation can be obtained:

$$\nabla_{\mathbf{w}}\mathrm{G}(n) = -2E[e_n^*\mathbf{x}_n] + \mathbf{C}\boldsymbol{\gamma}, \tag{8}$$

where $\nabla_{\mathbf{w}}\mathrm{G}(n)$ denotes the gradient vector.

In the real-time estimation, the equation (8) can be written as:

$$\nabla_{\mathbf{w}}\mathrm{G}(n) = -2e_n^*\mathbf{x}_n + \mathbf{C}\boldsymbol{\gamma}. \tag{9}$$

Using the gradient descent principle, the update equation in the objective function is expressed as:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \frac{\mu}{2}\nabla_{\mathbf{w}}\mathrm{G}(n). \tag{10}$$

Substituting (8) into (10), we can get the final update equation:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e_n^*\mathbf{x}_n - \frac{\mu}{2}\mathbf{C}\boldsymbol{\gamma}. \tag{11}$$

Multiply both sides of (11) by $\mathbf{C}^{\mathrm{H}}$, and combine (6) to find $\boldsymbol{\gamma}$:

$$\boldsymbol{\gamma} = (\mathbf{C}^{\mathrm{H}}\mathbf{C})^{-1}\left(\frac{2}{\mu}\mathbf{C}^{\mathrm{H}}\mathbf{w}_n + 2e_n^*\mathbf{C}^{\mathrm{H}}\mathbf{x}_n - \frac{2}{\mu}\mathbf{F}\right). \tag{12}$$

Re-substituting the obtained result (12) into (11) gives the updated equation:

$$\mathbf{w}_{n+1} = \mathbf{P}(\mathbf{w}_n + \mu e_n^*\mathbf{x}_n) + \mathbf{F}_{\mathbf{z}}. \tag{13}$$

In equation (13), we have:

$$\mathbf{P} = \mathbf{I}_{M\times M} - \mathbf{C}(\mathbf{C}^{\mathrm{H}}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{H}}, \tag{14}$$

$$\mathbf{F}_{\mathbf{z}} = \mathbf{C}(\mathbf{C}^{\mathrm{H}}\mathbf{C})^{-1}\mathbf{F}. \tag{15}$$

B. The CNLMS algorithm

In the CLMS algorithm, the value of the step size μ is fixed. In order to speed up the convergence of the CLMS algorithm, the square of the posterior a posteriori error with respect to the step size at the snapshot point n can be minimized [4], [23]:

$$\frac{\partial\left[|e_{ap}(n)|^2\right]}{\partial u_n^*} = \frac{\partial[e_{ap}(n)e_{ap}^*(n)]}{\partial \mu_n^*} = 0, \tag{16}$$

where $e_{ap}(n) = d_n - \mathbf{w}_{n+1}^{\mathrm{H}}\mathbf{x}_n = e_n(1 - u_n\mathbf{x}_n^{\mathrm{H}}\mathbf{P}\mathbf{x}_n)$.

Solving (16), the step-size at the nth snapshot point can be obtained:

$$\mu_n = \frac{\mu_0}{\mathbf{x}_n^{\mathrm{H}}\mathbf{P}\mathbf{x}_n + \beta}, \tag{17}$$

where μ_0 is a fixed step size, which is the initial value of the algorithm's step size, and β is positive that is close to zero which is to prevent the denominator from being zero in (17).

After the previous calculations, the update equation of the CNLMS algorithm is as:

$$\mathbf{w}_{n+1} = \mathbf{P}\left[\mathbf{w}_n + \mu_0\frac{e_n^*\mathbf{x}_n}{\mathbf{x}_n^{\mathrm{H}}\mathbf{P}\mathbf{x}_n + \beta}\right] + \mathbf{F}_{\mathbf{z}}. \tag{18}$$

IV. THE PROPOSED CIM-CNLMS ALGORITHM

In our proposed algorithm, the CIM is considered to approximate the l_0-norm to create a new cost function based on entropy theory [24-26]. As we know, the similarity between two vectors **X** and **Y** is measured using correntropy theory. The mathematical definition of the correlation entropy is presented as:

$$\mathrm{V}(\mathbf{X},\mathbf{Y}) = \frac{1}{M}\sum_{i=1}^{M} k(x_i, y_i), \tag{19}$$

where $\mathbf{X} = [x_1, x_2, \ldots, x_M]$, $\mathbf{Y} = [y_1, y_2, \ldots, y_M]$,and $k(.)$ is a regenerative kernel function. In the equation, the Gaussian kernel function is given by:

$$k(x,y) = k(x-y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right), \tag{20}$$

where σ is the kernel width. The l_0-norm of an M-dimensional vector $\mathbf{w} = [\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_M]$ is defined mathematically as:

$$\|\mathbf{w}\|_0 = \mathrm{card}\{w_i : w_i \neq 0\}, \tag{21}$$

where the l_0-norm of the vector $\mathbf{w}$ is the number of non-zero entries in $\mathbf{w}$ and card is set the cardinality [27].

As we know, solving l_0-norm is an NP hard problem [27]. Hence, continuous function is often utilized to approximate l_0-norm [28]. In our proposed algorithm, the CIM theory is used to approximate the l_0-norm to further develop the sparsity of the CNLMS algorithm. Therefore, our approximation l_0-norm approximation:

$$\begin{aligned}\|\mathbf{w}\|_0 \sim \mathrm{CIM}(\mathbf{w},0) &= \sqrt{\mathrm{k}(0) - \frac{1}{\mathrm{M}}\sum_{\mathrm{i}=1}^{\mathrm{M}}\mathrm{k}(\mathrm{w}_i,0)} \\ &= \sqrt{\frac{\mathrm{k}(0)}{\mathrm{M}}\sum_{i=1}^{\mathrm{M}}\left\{1 - \exp\left(-\frac{\mathrm{w}_i^2}{2\sigma^2}\right)\right\}}.\end{aligned} \tag{22}$$

To simplify the expression in (22), the square root operation in (22) is removed, resulting in the approximation:

$$\|\mathbf{w}\|_0 \sim \mathrm{CIM}^2(\mathbf{w},0) = \frac{\mathrm{k}(0)}{\mathrm{M}}\sum_{i=1}^{\mathrm{M}}\left\{1 - \exp\left(-\frac{\mathrm{w}_i^2}{2\sigma^2}\right)\right\}, \tag{23}$$

with $k(0) = \frac{1}{\sqrt{2\pi}\sigma}$.

Then, a l_0-norm penalty on the filter coefficients is integrated in the cost function of the LMS algorithm to accelerate coefficient convergence. In our proposed algorithm, a new l_0-CNLMS algorithm is developed for adaptive beamforming control to make the coefficients toward sparseness:

$$\min_{\mathbf{w}} E[|\mathrm{e}_n|^2] \ \text{s.t.} \begin{cases} \mathbf{C}^{\mathrm{H}}\mathbf{w} = \mathbf{F}, \\ \|\mathbf{w}\|_0 = t, \end{cases} \tag{24}$$

where $\|\mathbf{w}\|_0$ represents the number of non-zero entries in $\mathbf{w}$, and t represents the constraint value of the l_0-norm of $\mathbf{w}$.

The cost function with l_0-norm penalty is defined as:

$$\mathrm{G}_n^{l_0} = \mathrm{E}[|\mathrm{e}_n|^2] + \boldsymbol{\gamma}^{\mathrm{H}}(\mathbf{C}^{\mathrm{H}}\mathbf{w} - \mathbf{F}) + \gamma_1(\|\mathbf{w}\|_0 - \mathrm{t}). \tag{25}$$

According to (23), we can rewrite (25) as:

$$\begin{aligned} \mathrm{G}_n^{l_0} &= \mathrm{E}[|\mathrm{e}_n|^2] + \boldsymbol{\gamma}^{\mathrm{H}}(\mathbf{C}^{\mathrm{H}}\mathbf{w} - \mathbf{F}) \\ &+\gamma_1 \left\{ \frac{1}{\mathrm{M}\sqrt{2\pi}\sigma} \sum_{i=1}^{\mathrm{M}} \left[1 - \exp\left(-\frac{\mathrm{w}_i^2}{2\sigma^2} \right) \right] - t \right\}. \end{aligned} \tag{26}$$

Then, we have:

$$\begin{cases} \nabla_{\mathbf{w}} \mathrm{G}_n^{l_0} = -2\mathrm{e}_n^* \mathbf{x}_n + \mathbf{C}\boldsymbol{\gamma} + \gamma_1 \mathbf{J}_{\mathrm{n}}, \\ \mathbf{J}_n = \dfrac{1}{\mathrm{M}\sqrt{2\pi}\sigma^3} \left[\mathrm{w}_1 \exp\left(\dfrac{-{\mathrm{w}_1}^2}{2\sigma^2} \right), \ldots, \mathrm{w}_M \exp\left(\dfrac{-{\mathrm{w}_M}^2}{2\sigma^2} \right) \right]^{\mathrm{H}}. \end{cases} \tag{27}$$

According to the gradient descent principle, the weight coefficient update equation of the CIM-CNLMS algorithm can be obtained:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \frac{\mu}{2}\{-2\mathrm{e}_n^* \mathbf{x}_n + \mathbf{C}\boldsymbol{\gamma} + \gamma_1 \mathbf{J}_n\}. \tag{28}$$

Next, the coefficient $\boldsymbol{\gamma}$ is gotten by multiplying $\mathbf{C}^{\mathrm{H}}$ at both sides of (28). According to the constraint condition, $\mathbf{C}^{\mathrm{H}}\mathbf{w}_n = \mathbf{C}^{\mathrm{H}}\mathbf{w}_{n+1} = \mathbf{F}$, we get,

$$\boldsymbol{\gamma} = (\mathbf{C}^{\mathrm{H}}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{H}}(2\mathrm{e}_n^* \mathbf{x}_n - \gamma_1 \mathbf{J}_n). \tag{29}$$

Here, we assume that the update equation in the algorithm has tended to converge, i.e., $\mathbf{w}_n = \mathbf{w}_{n+1}$. The approximation condition $\mathbf{J}_n^{\mathrm{H}}\mathbf{w}_{n+1} = \mathrm{t}$ is proposed in [8] because $\mathbf{w}_n$ and $\mathbf{w}_{n+1}$ are in the same quadrant. So, we have the constraints:

$$\mathbf{J}_n^{\mathrm{H}}\mathbf{w}_{n+1} = \mathrm{t}, \quad \mathbf{J}_n^{\mathrm{H}}\mathbf{w}_n = \mathrm{t}_n. \tag{30}$$

By multiplying the left and right sides of the equation in (28) by $\mathbf{J}_{\mathrm{n}}^{\mathrm{H}}$, the equation is obtained:

$$\mathrm{t} = \mathrm{t}_n - \frac{\mu}{2}\{-2e_n^* \mathbf{J}_n^{\mathrm{H}}\mathbf{x}_n + \mathbf{J}_n^{\mathrm{H}}\mathbf{C}\boldsymbol{\gamma} + \gamma_1 \mathbf{J}_n^{\mathrm{H}}\mathbf{J}_n\}. \tag{31}$$

Substituting the results in (29) into equation (31) and separating γ_1, we have:

$$\gamma_1 = -\frac{2}{\mu \mathrm{r}} e_0(n) + \frac{2e_n^* \mathbf{J}_n^{\mathrm{H}}\mathbf{P}\mathbf{x}_n}{\mathrm{r}}, \tag{32}$$

where $\mathrm{r} = \mathbf{J}_n^{\mathrm{H}}\mathbf{P}\mathbf{J}_n$ is a scalar, and $e_0(n) = \mathrm{t} - \mathrm{t}_n$.

Substituting (29) and (32) into (28), the final update equation is obtained:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \frac{e_0(n)}{\mathrm{r}}\mathbf{P}\mathbf{J}_n + \mu e_n^* \left(\mathbf{P}\mathbf{x}_n - \frac{\mathbf{J}_n^{\mathrm{H}}\mathbf{P}\mathbf{x}_n \mathbf{P}\mathbf{J}_n}{\mathrm{r}} \right). \tag{33}$$

We set:

$$\begin{cases} \mathrm{r} = \mathbf{J}_n^{\mathrm{H}}\mathbf{P}\mathbf{J}_n, \\ e_0(n) = \mathrm{t} - \mathbf{J}_n^{\mathrm{H}}\mathbf{w}_n, \\ \mathbf{F}_0(n) = \dfrac{\mathrm{e}_0(n)}{\mathrm{r}}\mathbf{P}\mathbf{J}_n, \\ \mathrm{c} = \dfrac{\mathbf{J}_n^{\mathrm{H}}\mathbf{P}\mathbf{x}_n}{\mathrm{r}}. \end{cases} \tag{34}$$

Then (33) can be written as:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e_n^* \mathbf{P}(\mathbf{x}_n - \mathrm{c}\mathbf{J}_n) + \mathbf{F}_0(n). \tag{35}$$

The previously derived normalized version of the CLMS algorithm is utilized in the l_0-CLMS algorithm and l_0-norm is approximated using CIM. Substituting (35) into $e_{ap}(n) = d_n - \mathbf{w}_{n+1}^{\mathrm{H}}\mathbf{x}_n$ gives the equation:

$$e_{ap}(n) = e_n[1 - \mu_n(\mathbf{x}_n^{\mathrm{H}} - \mathrm{c}^* \mathbf{J}_n^{\mathrm{H}})\mathbf{P}\mathbf{x}_n] - \mathbf{F}_0^{\mathrm{H}}(n)\mathbf{x}_n. \tag{36}$$

Based on the previous derivation of the CNLMS algorithm, the step size of the new CIM -CNLMS algorithm is realized and given by:

$$\mu_n = \frac{\mu_0[e_n - \mathbf{F}_0^{\mathrm{H}}(n)\mathbf{x}_n]}{e_n(\mathbf{x}_n^{\mathrm{H}} - \mathrm{c}^* \mathbf{J}_n^{\mathrm{H}})\mathbf{P}\mathbf{x}_n + \alpha}. \tag{37}$$

A fixed convergence control factor μ_0 is introduced to control the offset, and α is a positive parameter that is close to 0 to prevent the denominator from being 0 in the equation (37). By substituting (37) into (35), the updating equation of the new CIM-CNLMS is obtained and presented as:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_n \mathrm{e}_n^* \mathbf{P}(\mathbf{x}_n - \mathrm{c}\mathbf{J}_n) + \mathbf{F}_0(n), \tag{38}$$

where

$$\begin{cases} \mathbf{P} = \mathbf{I}_{M\times M} - \mathbf{C}(\mathbf{C}^{\mathrm{H}}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{H}}, \\ \mathrm{c} = \dfrac{\mathbf{J}_n^{\mathrm{H}}\mathbf{P}\mathbf{x}_n}{\mathrm{r}}, \\ e_n = d_n - \mathbf{w}^{\mathrm{H}}\mathbf{x}_n, \\ \mu_n = \dfrac{\mu_0[\mathrm{e}_n - \mathbf{F}_0^{\mathrm{H}}(n)\mathbf{x}_n]}{e_n(\mathbf{x}_n^{\mathrm{H}} - \mathrm{c}^* \mathbf{J}_n^{\mathrm{H}})\mathbf{P}\mathbf{x}_n + \alpha}, \\ \mathrm{r} = \mathbf{J}_n^{\mathrm{H}}\mathbf{P}\mathbf{J}_n, \\ \mathbf{F}_0(n) = \dfrac{e_0(n)}{\mathrm{r}}\mathbf{P}\mathbf{J}_n. \end{cases} \tag{39}$$

V. SIMULATION RESULTS

In this section, we constructed several simulation experiments to test the performance of the proposed algorithm, whose performance is compared with the algorithms mentioned in the previous references [18]. The signals used in the simulation experiment are QPSK signals, four interference signals and one SOI signal. The horizontal azimuth of the SOI signal is 90°, and the horizontal azimuth of the four interference signals are respectively 80°, 22°, 52°, 147°. The zeniths of all the signals are 45°. The signal-to-noise ratio (SNR) is set to 20 dB, and the interference-to-noise ratio (INR) is set to 40 dB. The frequency for all the experiments is 2 GHz. These signals are received by a regular hexagonal array (HA) with 91 elements. In the simulation, the initial steps μ_0 for the CIM-CNLMS algorithm, the CNLMS algorithm, and the l_1-WCNLMS algorithm are 8×10^{-2}, 5×10^{-3}, 5×10^{-2}, respectively. In the simulation experiment, the constraint factor t of the CIM-CNLMS algorithm is 0.4, and the constraint factor t of the l_1 -WCNLMS algorithm is 1.07. In the simulation experiment, the number of iterations of the data is 12000. In the experiment, $\sigma = 0.0032$ is used in the CIM-CNLMS algorithm.

Figure 2 shows the beam pattern performance of

the proposed algorithm, and compares the beam pattern performance of the proposed algorithm with other existing algorithms. Our proposed CIM-CNLMS algorithm forms a null at the horizontal incident angle of the interfering signal, providing almost identical main lobe at the incident direction of the SOI. In addition, we can clearly see in the figure that the proposed algorithm has lower side lobe level (SLL) compared to l_1-WCNLMS. Therefore, the proposed algorithm has better beam performance than the l_1-WCNLMS algorithm.

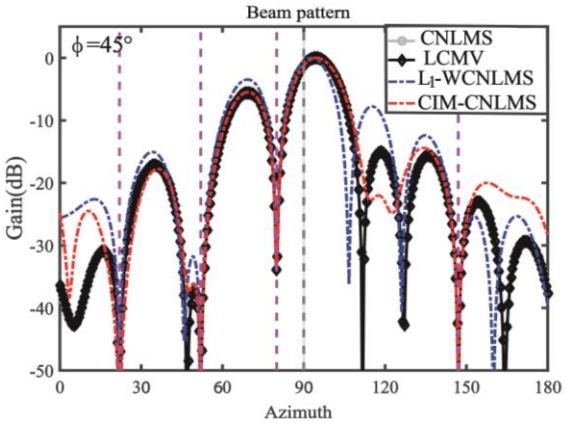
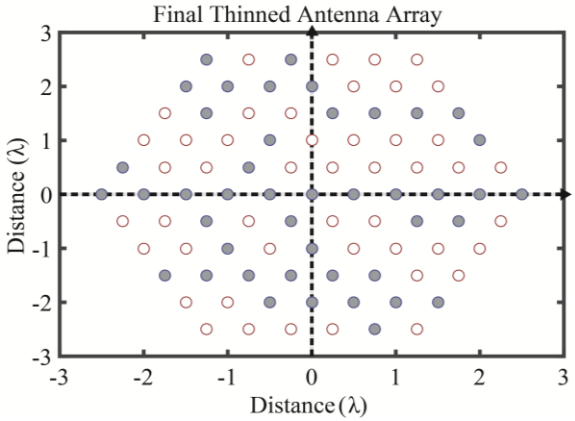


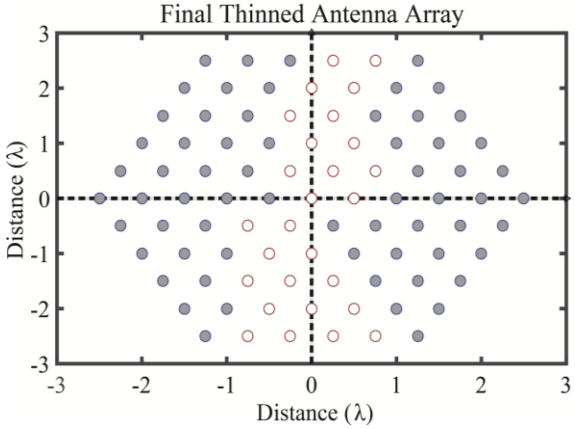
Fig. 2. Beam patterns comparison of CIM-CNLMS algorithm with the LCMV, CNLMS and l_1-WCNLMS algorithms. The pink dot lines represent the interferences, and the black dot line represents the SOI.

Figure 3 (a) is a sparse array generated by proposed CIM-CNLMS algorithm. The number of active array elements is 44 in an array with 91 elements, resulting in the sparsity of our proposed algorithm of 48.4%. Figure 3 (b) is a sparse array generated by the l_1-WCNLMS algorithm. The number of active array elements is 64, leading to the sparseness of the sparse array of 70.3%. Comparing the two previous algorithms, we can clearly see that our proposed CIM-CNLMS algorithm is far superior to the l_1-WCNLMS algorithm in terms of the finalized sparsity and the performance.

Figure 4 shows the MSE of the three algorithms, where the blue line represents the MSE of the CNLMS algorithm, the red line represents the MSE of the proposed CIM-CNLMS algorithm, and the yellow line denotes the MSE of the l_1-WCNLMS algorithm. From the figure, the proposed algorithm has the same MSE value as the l_1-WCNLMS algorithm after convergence, but our proposed algorithm converges at 1000th iteration, while the l_1-WCNLMS algorithm converges at 4000th iteration, and the CNLMS algorithm converges at the 8000th iteration. Therefore, the proposed algorithm converges the fastest. It is not difficult to conclude that the proposed algorithm has the best MSE performance among these algorithms.



(a) Sparse array thinned by CIM-CNLMS algorithm



(b) Sparse array thinned by l_1-WCNLMS algorithm

Fig. 3. The thinned sparse arrays with black dots of active array elements and white dots of inactive array elements.

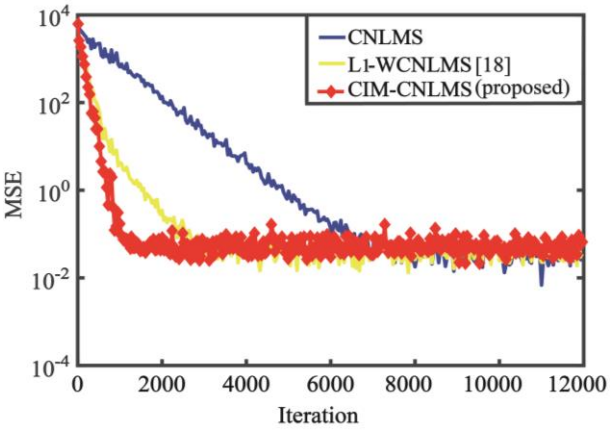


Fig. 4. Convergence of the proposed CIM-CNLMS algorithm.

The SINR results of the CIM-CNLMS are presented in Fig.5. From Fig.5, we can see that the SINR of the proposed CIM-CNLMS algorithm is better than the l_1-WCNLMS algorithm. However, it should be further

improved to get a nearly same SINR with the optimal one. The proposed algorithm can also be used for sparse DOA applications like [29] and uses the block norm in [30-33]. In addition, the proposed techniques can also be used in MIMO arrays [34-36] and UAV systems [37].

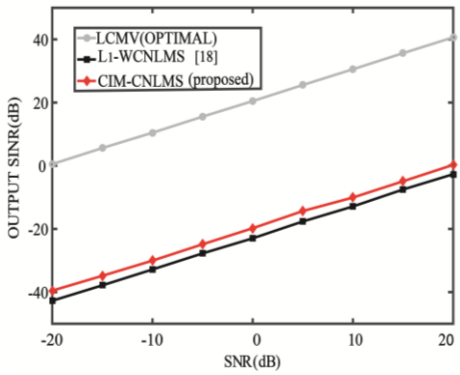


Fig. 5. SINR of the proposed CIM-CNLMS algorithm.

VI. CONCLUSION

In this paper, an improved adaptive beamforming algorithm, which is named as correntropy induced metric based constrained normalized least mean square (CIM-CNLMS), has been proposed and investigated for thinning arrays to reduce the computations and exploiting the sparsity. The CIM-CNLMS algorithm remained main lobe in the direction of SOI, and suppressed the interferences using nulls in the direction of the interferences. The simulation results demonstrated that the proposed CIM-CNLMS algorithm reduces the number of active array elements for achieving the desired performance like the l_1-WCNLMS algorithm. Additionally, the proposed algorithm has good beam pattern performance and better MSE performance in comparison with the popular algorithms mentioned in this paper. In addition, the output SINR of the proposed algorithm is better than the l_1-WCNLMS algorithm.

ACKNOWLEDGMENT

This research was funded by the National Natural Science Foundation of China (Grant No. 51779061, 51979061) and the Innovation Special Zone of National Defense Science and Technology.

REFERENCES

[1] J. Li and P. Stoica (Eds.), *Robust Adaptive Beamforming*. John Wiley & Sons, New York, NY, 2005.

[2] H. L. Van, *Trees, Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing.* John Wiley & Sons, New York, NY, 2002.

[3] O. L. Frost III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926-935, Aug. 1972.

[4] J. A. Apolinário, S. Werner, P. S. R. Diniz, and T. I. Laakso, "Constrained normalized adaptive filtering for CDMA mobile communications," *IEEE Signal Processing Conference*, Rhodes, Greece, pp. 1-4, Sep. 1998.

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

[6] Y. Chen, Y. Gu, and A. O. Hero, "Sparse LMS for system identification," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, pp. 3125-3128, Apr. 2009.

[7] O. Taheri and S. A. Vorobyov, "Sparse channel estimation with l_p-norm and reweighted l_1-norm penalized least mean squares," *IEEE International Conference on Acoustic Speech and Signal Processing (ICASS'11)*, Prague, Czech Republic, pp. 2864-2867, May 2011.

[8] Y. Li, Z. Jiang, O. M. Omer-Osman, Xiao Han, and Jingwei Yin, "Mixed norm constrained sparse APA algorithm for satellite and network echo channel estimation," *IEEE Access*, vol. 6, pp. 65901-65908, 2018.

[9] Y. Gu, J. Jin, and S. Mei, "l_0-norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 774-777, Sep. 2009.

[10] Y. Li, Y. Wang, R. Yang, et al., "A soft parameter function penalized normalized maximum correntropy criterion algorithm for sparse system identification," *Entropy*, vol. 19, no. 1, p. 45, Jan. 2017. 10.3390/ e19010045.

[11] Z. Jin, Y. Li, and J. Liu, "An improved set membership proportionate adaptive algorithm for a block-sparse system," *Symmetry*, vol. 10, no. 3, p. 75, 2018.

[12] D. Angelosante, J. A. Bazerque, and G. B. Giannakis, "Online adaptive estimation of sparse signals: Where RLS meets the *l*1-norm," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3436-3447, Mar. 2010.

[13] O. Taheri and S. A. Vorobyov, "Reweighted l_0-norm penalized LMS for sparse channel estimation and its analysis," *Signal Processing*, vol. 104, pp. 70-79, May 2014.

[14] Y. Li, Y. Wang, and T. Jiang, "Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation," *AEU - International Journal of Electronics and Communications*, vol. 70, no. 7, pp. 895-902, 2016.

[15] Y. Li, Y. Wang, and T. Jiang, "Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation," *Signal Processing*, 128:243-251, 2016.

[16] R. Tibshirani, "Regression shrinkage and selection

via the lasso," *J. R.Stat. Soc. Ser. B-Stat. Methodol.*, vol. 58, no. 1, pp. 267-288, Jan. 1996.

[17] J. F. de Andrade, Jr., M. L. R. de Campos, and J. A. Apolinário, Jr., "An l_1 -norm linearly constrained LMS algorithm applied to adaptive beamforming," in *Proc. 7th IEEE Sensor Array Multichannel Signal Process. Workshop (SAM)*, Hoboken, NJ, USA, pp. 429-432, June 2012.

[18] J. F. de Andrade, M. L. R. de Campos, and J. A. Apolinário, "l_1-constrained normalized LMS algorithms for adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6524-6539, Dec. 2015.

[19] W. Shi, Y. Li, and S. Luo, "Adaptive antenna array beamforming based on norm penalized NLMS algorithm," *IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, in press, Boston, America, July 2018.

[20] W. Shi and Y. Li, "Norm-constrained NLMS for sparse controllable adaptive array beamforming," *International Applied Computational Electromagnetics Society Symposium*, Beijing, China, July 2018.

[21] W. Shi, Y. Li, and J. Yin, "Improved constraint NLMS algorithm for sparse adaptive array beamforming control applications," *Applied Computational Electromagnetics Society Journal*, vol. 34, no. 3, pp. 419-424, 2019.

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

[23] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation, 3rd ed.*, New York, NY, USA: Springer, 2010.

[24] W. Ma, Q. Hua, G. Gui, et al., "Maximum correntropy criterion based sparse adaptive filtering algorithms for robust channel estimation under non-Gaussian environments," *Journal of the Franklin Institute*, vol. 352, no. 7, pp. 2708-2727, 2015.

[25] R. He, W. S. Zheng, and B. G Hu, "Maximum correntropy criterion for robust face recognition," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 33, no. 8, pp. 1561-1576, 2011.

[26] X. Xu and Y. B. Li, "A multiscale hierarchical threshold-based completed local entropy binary pattern for texture classification, cognitive computation," vol. 12, no. 1, pp. 224-237, 2020.

[27] J. Weston, A. Elisseeff, B. Schölkopf, and M. Tipping, "Use of zero-norm with linear models and kernel methods," *JMLR special Issue on Variable and Feature Selection*, pp. 1439-1461, 2002.

[28] S. Seth and J. Príncipe, "Compressed signal reconstruction using the correntropy induced metric," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 3845-3848, 2008.

[29] X. Zhang, T. Jiang, Y. Li, et al., "A novel block sparse reconstruction method for DOA estimation with unknown mutual coupling," *IEEE Communications Letters*, vol. 23, no. 10, pp. 1845-1848, 2019.

[30] Y. Li, Z. Jiang, W. Shi, X. Han, and B. D. Chen, "Blocked maximum correntropy criterion algorithm for cluster-sparse system identification," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 11, pp. 1915-1919, 2019.

[31] B. Chen, Z. Li, Y. Li, et al., "Asymmetric correntropy for robust adaptive filtering," arXiv preprint arXiv:1911.11855.

[32] Q. Wu, Y. Li, and Y. V. Zakharov, "A kernel affine projection-like algorithm in reproducing kernel Hilbert space," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 10.1109/TCSII.2019. 2947317.

[33] W. Shi, Y. Li, and B. Chen, "A separable maximum correntropy adaptive algorithm," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 10.1109/TCSII.2020.2977608.

[34] K. Yu, Y. Li, and X. Liu, "Mutual coupling reduction of a MIMO antenna array using 3-D novel meta-material structures," *Applied Computational Electromagnetics Society Journal*, vol. 33, no. 7, pp. 758-763, 2018.

[35] T. Jiang, T. Jiao, Y. Li, and W. Yu, "A low mutual coupling MIMO antenna using periodic multi-layered electromagnetic band gap structures," *Applied Computational Electromagnetics Society Journal*, vol. 33, no. 3, pp. 305-311, 2018.

[36] J. Jiang, Y. Xia, and Y. Li, "High isolated X-band MIMO array using novel wheel-like metamaterial decoupling structure," *Applied Computational Electromagnetics Society Journal*, vol. 34, no. 12, pp. 1829-1836, 2019.

[37] C. Su, F. Ye, L. Wang, et al., "UAV-Assisted wireless charging for energy-constrained IoT devices using dynamic matching," *IEEE Internet of Things Journal*, Doi: 10.1109/JIOT.2020.2968346.

Efficient Indoor Signal Propagation Model Based on LOLA-Voronoi Adaptive Meshing

Junyi Yao [1], Wanchun Tang [*2, 3], Baozhu Li [1], Shuming Zhang [1], and Rui Sun [1]

[1] School of Physics and Technology
Nanjing Normal University, Nanjing, 210023, China
1016424781@qq.com, 951408095@qq.com, 1627839409@qq.com, 1067311237@qq.com

[2] School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing, 210023, China

[3] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application
Nanjing, 210023, China
eewctang@njnu.edu.cn

Abstract ─ An innovative adaptive mesh strategy based on LOLA-Voronoi (Local Linear Approximation-Voronoi) is proposed to efficiently predict indoor signal propagation. This indoor high-efficiency propagation model (HPMO) can identify nonlinear regions to capture the complex behavior and achieve sufficient prediction accuracy when the computational cost is limited. A set of representative reference scenario simulation settings and results are reported and discussed to analyze the accuracy, and the efficiency of HPMO. Comparison with the original model based on traditional uniform mesh shows that the proposed method herein yields a considerable reduction in the prediction calculation cost of the complex indoor environment, while maintaining sufficient accuracy.

Index Terms ─ HPMO, indoor signal propagation model, LOLA-Voronoi, received signal strength.

I. INTRODUCTION

With the growing demand for indoor wireless LANs and personal communication networks, it is very important to adequately consider wireless communication in indoor environments. Recently, indoor localization [1-2], wireless communication system design [3], and human exposure assessment [4] have become popular application areas for indoor electromagnetic environments. Therefore, efficient prediction of indoor radio wave propagation becomes a necessary basic step.

A wide range of methods focused on indoor propagation prediction have been proposed. Deterministic models (such as FDTD, UTD, MoM [4-7]) have high accuracy, however, they are time-consuming due to the inherent computational complexity. On the contrary, empirical models (such as ITU-R, Motley-Keenan, Okumura–Hata, COST-231 [8-11]) have the advantage of rapid implementation, but the accuracy is flawed because the environment is simplified. In [12], multipath effect is considered in the model of radio channel, polarized channel model has also been defined within [13]. A modified Motley-Keenan model based on ray-tracing was developed in [14-15], which takes into account both the multipath effect and beam polarization, thus obtaining the trade-off between accuracy and efficiency. Unfortunately, just uniform meshing is considered in this model and with higher operation frequency, smaller cell sizes, and more advanced antenna systems, radio propagation modeling becomes more difficult and challenging.

To reduce the cost, a new zone-based propagation model based on reduced number of measurements to recover the indoor fingerprint database is taken in [16]. And a reconstruction technology based on 2D LOLA-Voronoi (Local Linear Approximation-Voronoi) adaptive meshing method without any prior knowledge on the source antenna is developed in [17].

LOLA-Voronoi used in [17] is an effective integration of a sequential experimental design (SED). Unlike fixed mesh, LOLA-Voronoi could use information gathered from previous data points to determine the position of new data points adaptively [18]. The outcome of LOLA-Voronoi is a representative set of data samples that is more concentrated within those regions in which larger deviations from a local linear approximation have been calculated starting from a small initial set of samples. And its practical application in complex indoor environment remains to be solved.

This paper aims at proposing an indoor high-efficiency propagation prediction method based on LOLA-Voronoi adaptive meshing. The main contributions are presented here: (a) it's the first attempt to implement an adaptive meshing using LOLA-Voronoi in an indoor

Submitted On: December 9, 2019
Accepted On: February 23, 2020

signal strength prediction developed by a modified Motley-Keenan model [14]; (b) the overall cost is reduced, while adequate accuracy is obtained in the computed estimation; (c) an indoor environment is selected as a reference scenario, and RMSE values of different indoor propagation models are calculated to verify the feasibility of the proposed model.

This paper is organized as follows. The modified indoor propagation model and formulation of the LOLA-Voronoi adaptive meshing method and the proposed HPMO are described in Section II. Section III describes the numerical simulation settings, prediction results and comparisons. Section IV concludes the paper.

II. HIGH-EFFICIENCY PROPAGATION MODEL IMPLEMENTATION

A. Modified Motley-Keenan model

The path loss of indoor empirical propagation models are generally expressed as follows:

$$L(d)=L(d_0)+20\log(d)+A\,, \tag{1}$$

where $L(d_0)$ is the free-space loss at a reference distance d_0 of 1m. d is the separation distance between transmit antenna (Tx) and receive antenna (Rx), A represents attenuation in signal strength, which tends to be variant in different models.

A modified Motley-Keenan model [9] based on ray tracing is adopted in this work, A in equation (1) is mathematically expressed as:

$$A=20\log(\frac{4\pi}{\lambda})-10\log(\prod\tau_w)-10\log(\prod\Gamma_w)\,, \tag{2}$$

where λ represents the wavelength, τ_k, Γ_k are respectively the transmission and reflection coefficient when the ray encounters obstacle k, which can be calculated by Fresnel coefficients [19].

The algorithm [15] uniformly divides the 2D research domain into a large number of nodes, and obtains the path loss corresponding to each node using the model expressed in equation (2).

More details are added to the received EM signal strength P (in dB), such as transmit power and antenna directivity:

$$P=P_{\mathrm{Tx}}+G_{\mathrm{Tx}}+G_{\mathrm{Rx}}-L(d)\,, \tag{3}$$

where P_{Tx} is the transmit power, G_{Tx} and G_{Rx} are the transmitter (Tx) and receiver (Rx) antenna gains respectively.

B. LOLA-Voronoi Adaptive meshing

Adaptive meshing based on LOLA-Voronoi is aimed at selecting nodes in unexplored regions of the design space and adding nodes in regions which were previously identified to be interesting such as nonlinear regions [18]. These two measures are combined into a hybrid score to calculate each existing node:

$$S(\mathbf{p}_m)=V(\mathbf{p}_m)+\frac{O(\mathbf{p}_m)}{\sum_{m=1}^{M}O(\mathbf{p}_m)}\,, \tag{4}$$

where $\mathbf{p}_m$ [m=1, 2, …, M] represents any one of the existing nodes, $V(\mathbf{p}_m)$ is the size of Voronoi cell where $\mathbf{p}_m$ is located, representing whether or not under-meshed around this point, the last fraction represents the non-linearity behavior at $\mathbf{p}_m$.

More specifically, Voronoi tessellation is an effective methodology to describe sampling density. It is a set of continuous polygons that connect the vertical bisectors of two adjacent line segments as shown in Fig. 1.

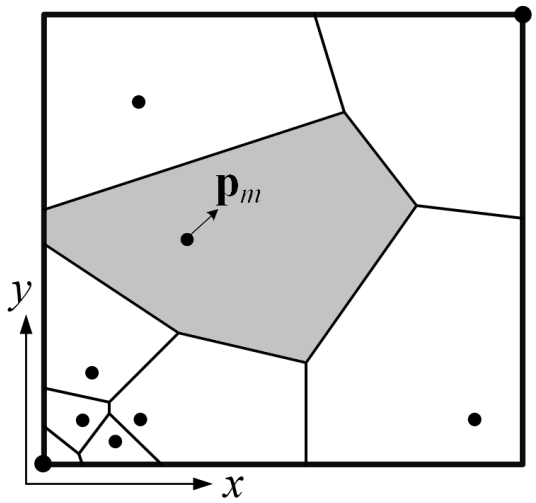


Fig. 1. Voronoi tessellation in design space.

Each cell contains only one node $\mathbf{p}_m$, and the cell size can be approximately estimated as:

$$V(\mathbf{p}_m)=\frac{k}{N}\,, \tag{5}$$

where N is the large number of randomly, evenly distributed Monte Carlo test points [$\mathbf{t}_n$, n=1, 2, …, N] generated in the domain, k denotes the number of test points closest to $\mathbf{p}_m$:

$$\left|\mathbf{t}_n-\mathbf{p}_m\right|\le\left|\mathbf{t}_n-\mathbf{p}_{m'}\right|\,|\,m\neq m'\,. \tag{6}$$

In addition, the non-linearity can be estimated by the gradient $\mathbf{g}$, which can be calculated by fitting a hyperplane through reference node $\mathbf{p}_m$ and its ith neighbor $\mathbf{p}_{mi}=(\mathbf{p}_{mi}^{(x)},\mathbf{p}_{mi}^{(y)})$ [i=1, 2, …, 4], the following system is solved by least squares:

$$\begin{pmatrix}\mathbf{p}_{m1}^{(x)}-\mathbf{p}_m^{(x)} & \mathbf{p}_{m1}^{(y)}-\mathbf{p}_m^{(y)}\\ \mathbf{p}_{m2}^{(x)}-\mathbf{p}_m^{(x)} & \mathbf{p}_{m2}^{(y)}-\mathbf{p}_m^{(y)}\\ \mathbf{p}_{m3}^{(x)}-\mathbf{p}_m^{(x)} & \mathbf{p}_{m3}^{(y)}-\mathbf{p}_m^{(y)}\\ \mathbf{p}_{m4}^{(x)}-\mathbf{p}_m^{(x)} & \mathbf{p}_{m4}^{(y)}-\mathbf{p}_m^{(y)}\end{pmatrix}\begin{pmatrix}\mathbf{g}_m^{(x)}\\ \mathbf{g}_m^{(y)}\end{pmatrix}=\begin{pmatrix}P(\mathbf{p}_{m1})\\ P(\mathbf{p}_{m2})\\ P(\mathbf{p}_{m3})\\ P(\mathbf{p}_{m4})\end{pmatrix}, \tag{7}$$

where $P(\mathbf{p}_{mi})$ is the output value (here specifically refers to the received signal strength) corresponding to $\mathbf{p}_{mi}$, $\mathbf{g}=(\mathbf{g}_m^{(x)},\mathbf{g}_m^{(y)})$ is the gradient that is being calculated.

Then, the nonlinearity of the system near $\mathbf{p}_m$ can be obtained by differences between the responses of the neighbors and the local linear approximation, using the following formula:

$$O(\mathbf{p}_m)=\sum_{i=1}^{4}\left|P(\mathbf{p}_{mi})-(P(\mathbf{p}_m)+\mathbf{g}\cdot(\mathbf{p}_{mi}-\mathbf{p}_m))\right|. \quad (8)$$

New node will be selected around the node with higher hybrid score according to equation (4).

C. High-efficiency propagation model based on LOLA-Voronoi

In order to avoid spending considerable amount of time and energy calculating a large number of nodes in [15], it is necessary to select a small number of interesting nodes to represent and predict the overall behavior of the propagation model.

Therefore, the work-flow of the proposed HPMO is illustrated in Fig. 2.

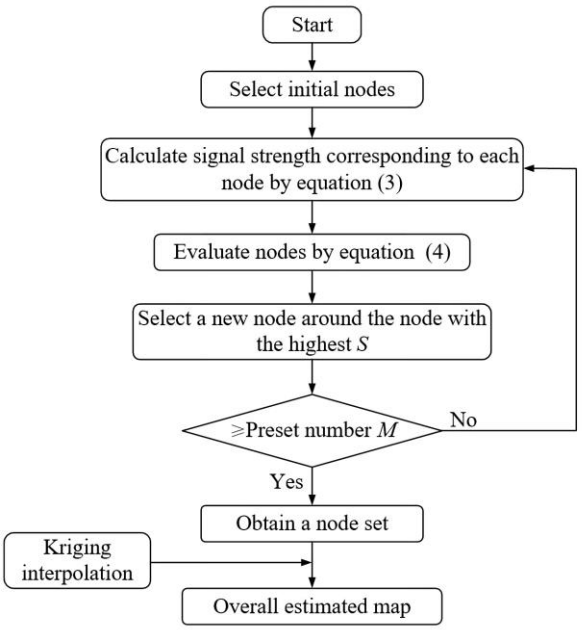


Fig. 2. Flow-chart of HPMO.

In the design process, from two initial nodes located on the diagonal of the two-dimensional space, their corresponding signal strength value are calculated by equation (3), a new node will be generated around the node with the highest ranked score S according to equation (4).

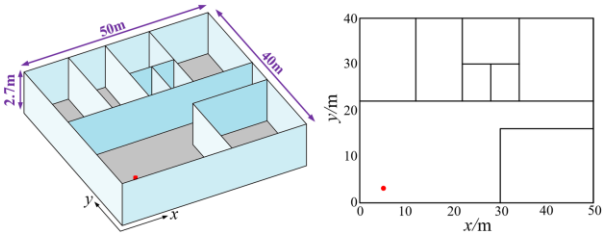
If the total number of the nodes has not yet been equal to the preset value M, a call is made to the model calculation and sequential design routine, which tends to select a new node to be evaluated, and the algorithm starts all over again. Finally, the EM signal strength distribution in the two-dimensional region is reconstructed according to the selected node and its corresponding output value.

III. NUMERICAL RESULTS

In this paper, all the simulation experiments are performed on a standard laptop with 1.6GHz CPU and 8GB of RAM.

A. Definition of indoor reference scenario

In order to gain insight into the proposed model, the considered environment in this paper is an empty indoor environment with a very complex topology and form, including the size and shape of different rooms and corridors as shown in Fig. 3.



(a) (b)

Fig. 3. Schematic view (a) and top view (b) of the indoor structure and transmitter location identification.

The dot in Fig. 3 corresponds to the transmitter antenna [dipole], which is located at coordinates (x = 5 m, y = 3 m, z = 0.2 m). Specific simulation parameters are defined in Table 1.

Table 1: Simulation parameters

Parameters	Value
Frequency	2.4GHz
Transmitter power	15dB(m)
Wall permittivity	5.3
Observation plane height	0.8m
Observation range(x)	[0m,50m]
Observation range(y)	[-5m,45m]

B. Uniform meshing

During the modeling process, in order to get better resolution, the uniform mesh size must be controlled below half the wavelength, which is due to Nyquist theory. As shown in Fig. 4 [mesh size of λ/2, total nodes of 6.4×10^5], the ray tracing algorithm [15] calculates each node to obtain a signal strength distribution map of the entire observation plane.

In this indoor structure example, the grid size is closely related to the calculation time, as shown in Fig. 5. The amount of computation for the exact solution of

the model tends to increase at an exponential rate with the problem size.

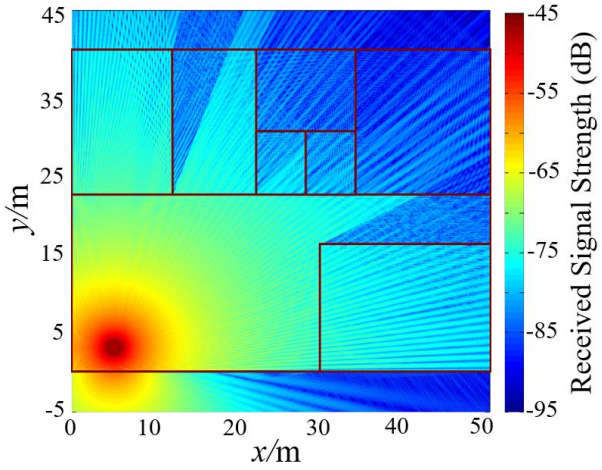


Fig. 4. EM signal strength propagation map obtained by modified Motley-Keenan model based on uniform meshing.

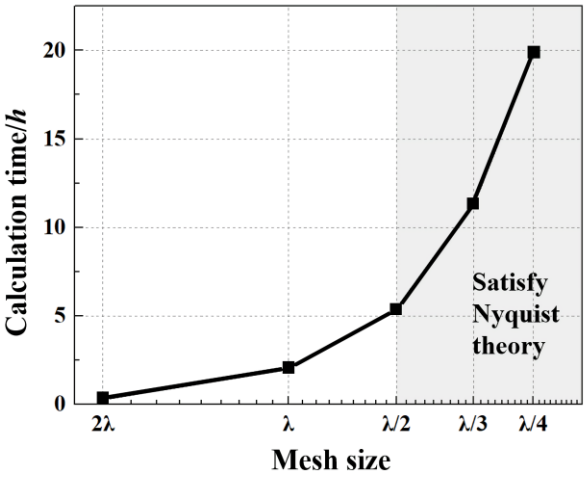


Fig. 5. The relation curve of mesh size and calculation time under uniform meshing.

C. Adaptive meshing used LOLA-Voronoi

In the study of electromagnetic field coverage in space, too fine mesh generation is not necessary. For the purpose of algorithm optimization and cost saving, the probes can be placed at interesting and representative positions, and then the signal strength in the entire 2D research area can be estimated by ordinary Kriging interpolation.

To illustrate how LOLA-Voronoi identifies nonlinear regions while still maintaining proper domain coverage, the node distributions have been obtained using the adaptive meshing strategy and random meshing respectively when the number of M = 200 [Fig. 6 (a), Fig. 7 (a)], and M = 500 [Fig. 6 (c), Fig. 7 (c)] nodes are used. Their corresponding prediction results are shown in Figs. 6 (b), (d) and Figs. 7 (b), (d).

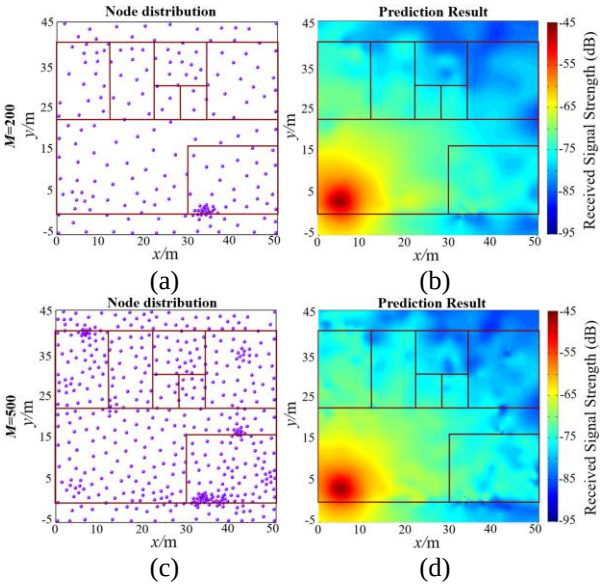


Fig. 6. The distribution of nodes by the LOLA-Voronoi adaptive meshing and the corresponding prediction map when (a), (b) M=200, (c), (d) M=500.

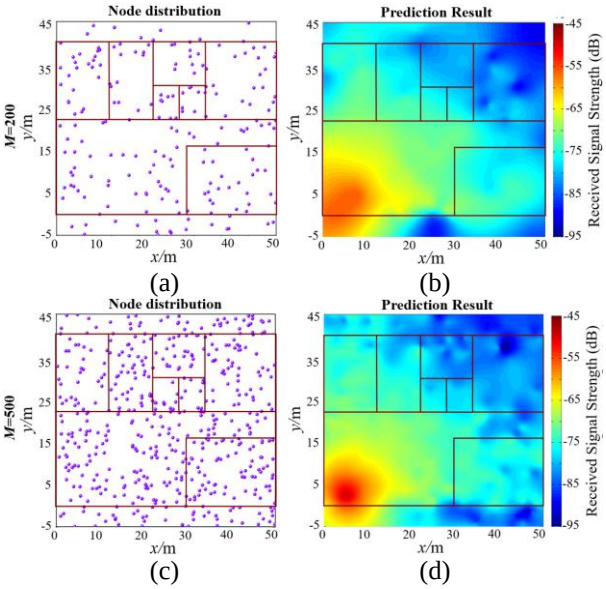


Fig. 7. The distribution of nodes by the random meshing and the corresponding prediction map when (a), (b) M=200, (c), (d) M=500.

As shown in Fig. 6, as the number of nodes increases, the intensive subdivision always occurs near the wall where the signal strength value is mutated, and under-meshed is also avoided in other areas. Because

LOLA-Voronoi is always able to capture the details of the changes in the field, and can adaptively identify the location of interest while maintaining proper coverage. Random meshing is selected for comparison, which usually lead to chaotic distribution of nodes and even under-meshed of some regions [Fig. 7].

For the intuitiveness and completeness of the conclusion, the HPMO based on LOLA-Voronoi adaptive meshing is further verified by the prediction error versus the number of nodes [Fig. 8, $10 \le M \le 1000$].

The prediction precision could be supported by defining the relative error (ξ):

$$\xi = \frac{\left\| \mathbf{s} - \hat{\mathbf{s}} \right\|_2}{\left\| \mathbf{s} \right\|_2}, \tag{9}$$

where $\mathbf{s}$ is the original data [Fig. 4] obtained by fine and uniform mesh, and $\hat{\mathbf{s}}$ is the estimated data based on the several meshing nodes. In this paper, all the prediction errors are obtained by taking the average of the relative errors of equation (9) after 50 times measurement.

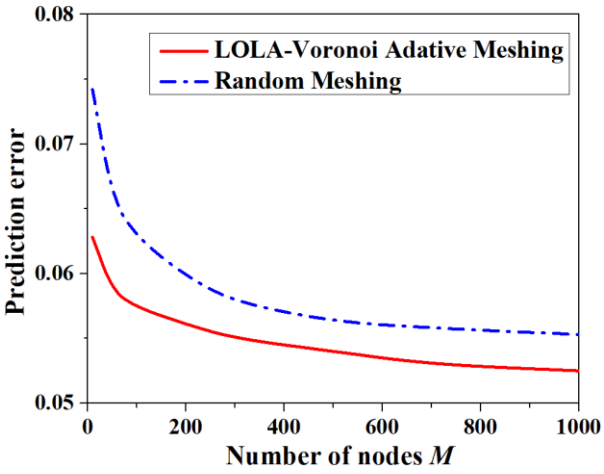


Fig. 8. *Performance Analysis*—Comparison of prediction error between LOLA-Voronoi adaptive meshing and random meshing.

In Fig. 8, LOLA-Voronoi adaptive meshing yields a certain improvement in the accuracy compared with random meshing. It is worth noticing that the fewer nodes makes the adaptive meshing more efficient than the random meshing, so as expected, the adaptive meshing strategy greatly accelerates the prediction of indoor propagation and keeps low prediction error by placing nodes in more important and interest positions.

In order to further prove the accuracy of the proposed HPMO, Table 2 lists the RMSE (Root Mean Square Error) of indoor propagation models based on the benchmark simulation data of commercial software Altair Winprop in the same reference scenario.

Table 2: RMSE of indoor propagation models

Models	RMSE(dB)
Log-distance	7.18
ITU-R [8]	8.46
Multi-wall [9],[11]	4.74
Modified Motley-Keenan [14-15]	4.65
HPMO (0.16% of total nodes)	4.69

It can be concluded from Table 2 that the HPMO proposed in this paper only needs 0.16% of the total number of nodes in the modified Motley-Keenan [14-15] model to reconstruct a result with the RMSE value of 4.69dB. It can be seen that even if the computing resources required by the original algorithm are reduced, it can still obtain sufficient prediction accuracy.

IV. CONCLUSION

In this paper, an innovative strategy for efficiently indoor propagation prediction has been proposed. In order to reduce the cost of the model program, the traditional dense uniform mesh method is replaced by an adaptive mesh method based on the output value. Accurately approximate the actual model based on a small number of interesting data points that are densely distributed in the nonlinear region and ensure that other regions are not under-meshed. Subsequently, the results from a representative set of numerical experiments have been reported and discussed to analyze the advantages of the proposed model in terms of accuracy and efficiency. The model obtained satisfactory results in the high-efficient modeling of indoor scene, and also provided an important insight into the position of measuring equipment in actual experimental measurement.

ACKNOWLEDGMENT

This work is supported by the National Key Research and Development Program of China (under Grant No. 2017YFB0503500) and the National Natural Science Foundation of China (under Grant No. 61571232).

The authors wish to thank Professor Salaheddin Hosseinzadeh for sharing the indoor radio propagation code online.

REFERENCES

[1] A. Achroufene, Y. Amirat, and A. Chibani, “RSS-based indoor localization using belief function theory,” *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 3, pp. 1163-1180, July 2019.

[2] S. Sun, Y. Li, W. S. T. Rowe, et al., “Practical evaluation of a crowdsourcing indoor localization system using hidden Markov models,” *IEEE Sensors Journal*, vol. 19, no. 20, pp. 9332-9340, Oct. 2019.

[3] S. Y. Lim, Z. Yun, and M. F. Iskander, “Radio

propagation modeling: A unified view of the ray-tracing image method across emerging indoor and outdoor environments," *The World of Applied Electromagnetics*, 2018.

[4] H. Y. Chen and S. H. Wen, "Evaluation of E-field distribution and human exposure for a LTE femtocell in an office," *Applied Computational Electromagnetics Society Journal (ACES)*, vol. 31, no. 4, pp. 455-467, Apr. 2016.

[5] J. Sosa, S. Coss, A. RodríGuez, et al., "Indoor 2.4GHz microwave propagation study using 3D FDTD approach," *Electronics Letters*, vol. 47, pp. 1308-1309, 2011.

[6] Z. Gu and Z. Jiang, *A Modified UTD Model for Indoor Propagation Prediction*. 2005.

[7] S. A. Mitilineos, S. C. Panagiotou, P. K. Varlamos, et al., "Indoor environments propagation simulation using a hybrid MoM and UTD electromagnetic method," *Annales des Telecommunications*, vol. 60, no. 9, pp. 1231-1243, 2005.

[8] "Propagation data and prediction methods for the planning of indoor radio communication systems and the radio local area networks in the frequency range 900 mhz to 100 ghz," *Recommendation ITU-R P.1238-8*, 2015.

[9] A. J. Motley and J. M. Keenan, "Radio coverage in buildings," *British Telecom Tech. J.*, vol. 8, no. 1, pp. 19-24, 1990.

[10] M. Hatay, "Empirical formula for propagation loss in land mobile radio services," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317-325, 1980.

[11] COST Action 231: Digital Mobile Radio Towards Future Generation Systems: Final Report, 1999.

[12] N. Iqbal, J. Luo, R. Müller, et al., "Multipath cluster fading statistics and modeling in millimeter-wave radio channels," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 4, pp. 2622-2632, Apr. 2019.

[13] R. Adeogun, T. Pedersen, C. Gustafson, et al., "Polarimetric wireless indoor channel modeling based on propagation graph," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 10, pp. 6585-6595, Oct. 2019.

[14] S. Hosseinzadeh, H. Larijani, K. Curtis, et al., "Empirical propagation performance evaluation of LoRa for indoor environment," *IEEE International Conference on Industrial Informatics*, 2017.

[15] S. Hosseinzadeh, "3D ray tracing for indoor radio propagation." (https: //www. mathworks.com/ matlabcentral/fileexchange/64695-3d-ray-tracing-for-indoor-radio-propagation) MATLAB Central File Exchange. Retrieved November 8, 2019.

[16] V. Moghtadaiee, S. A. Ghorashi, and M. Ghavami. "New reconstructed database for cost reduction in indoor fingerprinting localization," *IEEE Access*, vol. 7, pp. 104462-104477, 2019.

[17] B. Z. Li, M. Salucci, P. Rocca, J. Y. Yao, et al., "Total-variation compressive sensing based on hybrid sequential experimental design for field reconstruction," *Applied Computational Electromagnetics Society (ACES) Symposium*, Nanjing, China, Aug. 2019.

[18] K. Crombecq, D. Gorissen, D. Deschrijver, and T. Dhaene, "A novel hybrid sequential design strategy for global surrogate modeling of computer experiments," *SIAM Journal on Scientific Computing*, vol. 33, no. 4, pp. 1948-1974, Jan. 2011.

[19] "Effects of building materials and structures on radiowave propagation above about 100 MHz," *ITU-R P.2040-1*, 2015.

Junyi Yao was born in Jiangsu Province, China, in 1995. She received the B.S. degree from Nanjing Normal University (NJNU), Nanjing, in School of Physics and Technology in 2017. She is currently pursuing the M.S. degree with the School of Physics and Technology, NJNU.

Her current research interests include numerical calculation of EM field and reconstruction algorithm.

Wanchun Tang (M'04) was born in China in 1967. He received the B.S. degree from Tsinghua University, Beijing, China, in 1990, the M.S. degree from the Nanjing University of Science and Technology (NJUST), Nanjing, China, in 1995, and the Ph.D. degree from the City University of Hong Kong, Hong Kong, in 2003, all in Electrical Engineering.

He was a Full Professor with the Department of Communication Engineering, NJUST, and is currently a Specially Invited Full Professor with the Jiangsu Key Laboratory on Optoelectronic Technology, School of Physics and Technology, Nanjing Normal University, Nanjing. He has authored or coauthored over 100 journal and conference papers. His current research interests include modeling and optimization of RFIC, antennas, signal integrity, and power integrity design in package.

Dual Band, Miniaturized, Implantable Antenna Design with On-body Antennas for Wireless Health Monitoring

Ademola O. Kaka [1], Mehmet Toycan [1], Stuart D. Walker [2], and Doğa Kavaz [3]

[1] Cyprus International University, Faculty of Engineering, Department of Electric and Electronic Engineering
Nicosia, Northern Cyprus, Mersin 10, Turkey
mtoycan@ciu.edu.tr

[2] School of Computer Science and Electronic Engineering, University of Essex, Colchester, UK

[3] Cyprus International University, Faculty of Engineering, Bioengineering Department
Nicosia, Northern Cyprus, Mersin 10, Turkey
dkavaz@ciu.edu.tr

Abstract ─ A modified Hilbert fractal geometry and serpentine radiator-based implant antenna is proposed for dual band medical operations in both the MICS band (402-405 MHz) and ISM band (2.4–2.48 GHz). The antenna has miniaturized dimensions of 5.5 x 7.6 x 0.8 mm^3 (width x height x thickness) and is simulated inside a four layer (air, skin, fat and muscle) model with dimensions of 150 mm x 75mm x 55 mm. A Zirconia (ε_r=29) superstrate and shorting pin were utilized to achieve biocompatibility and the desired resonance. Relatively stable and omni-directional radiation patterns were achieved for the dual band operation. An ISM band on-body antenna (5.8 x 5.5 x 0.8 mm^3) was proposed to generate a wake-up signal while the wireless telemetry transmission was achieved using the MICS band on-body antenna (5.8 x 5.5 x 0.8 mm^3). Implant and on-body antennas demonstrated a reflection coefficient (S_{11}) better than -10 dB characteristics. To adhere to the SAR regulation limit of 1.6W/kg, the peak incident power should not exceed 0.25 mW and 0.2 mW for the ISM band and MISC band on-body antennas respectively. Propagation channel characteristics were simulated by observing the S_{12} and S_{13} characteristics and satisfactory results were achieved. The peak gain for the implant, MICS band and ISM band on-body antennas were -39.8 dBi, -35.6 dBi and -23 dBi, respectively due to the miniature dimensions. The miniaturized characteristics, dual-band operation, biocompatibility and stable characteristics in the presence of human tissue model make both the implant and on-body antennas suitable for biomedical monitoring systems.

Index Terms ─ Biocompatible, dual-band, implant antenna, miniaturized, on-body antenna.

I. INTRODUCTION

Wireless Body Area Network systems (WBANs) have great potential in a wide range of applications including health-monitoring, mobile-health, military etc., for the collection of real-time and stored physiological data [1,2]. Communication channels of WBANs are typically categorized as in-, on-, and off-body [3]. Implant antennas have a vital role for the provision of a high-quality communication link between the in-body and on-body medical units [4-8]. Compared to well-known antenna design techniques, implantable antennas should be miniaturized for physiological acceptability. In addition, these implants have some unique challenges with regards to impedance matching, bandwidth, multiband operation, low-power requirements (specific absorption rate limitation) and biocompatibility due to the variation in the electrical properties of human body tissues [6]. As demonstrated in previous studies, antenna characteristics must be optimized to achieve acceptable performance by accounting for detuning effects inside the human body which may otherwise have a considerable impact on the antenna response [9]. The implant antenna should possess satisfactory transmission characteristics in order to function as the communication link between an implanted device and an exterior instrument for biotelemetry. The Medical Implant Communications Service (MICS) band (402–405 MHz) is mainly used as the transmission channel between the implant unit and the on- or off-body units and is regulated by the United States Federal Communications Commission (1999) and the European Radio Communications Committee (1997). There is also an opportunity to operate in the Industrial, Scientific and Medical (ISM) bands of 915, 2450 and 5800 MHz

Submitted On: December 13, 2019
Accepted On: April 12, 2020

(Federal Communications Commission, 2012) to activate the implant medical unit [10]. An implant device may be switched between sleep and active modes to realize improve efficiency and extend the battery life of the device [11]. A raft of simulations for one- or three-layer tissue models have been proposed for the realization of in-depth analysis for surrounding complicated tissue environment [9,12] and subsequent experiments have investigated tissue-equivalent liquids [5,7], mimicking gels [11], and animals [13]. Several antenna designs have also been proposed with broadband and dual-band characteristics [14-18].

Antennas with dual band operation, in the MICS and ISM frequency bands, are achieved using Serpentine curves in [11-13] which have dimensions of 22.5 x 22.5 x 2.5 mm^3. Another antenna with smaller dimensions, 15Ã—22.5Ã—2.5 mm^3, can be used with a multilayer configuration and includes an independent line to feed radiating elements via electromagnetic coupling [12]. While this approach allows for greater miniaturization, the bandwidth requires further improvement. The Sierpinski fractal concept was utilized for Planar Inverted-F Antenna (PIFA) to achieve dual-band transmission [16] and a dual-band textile antenna based on Half-Mode Substrate Integrated Waveguide (HMSIW) in has been realized, albeit with a large form factor [18]. Circularly polarized (CP) radiation is usually preferred to linearly polarized radiation for wireless telemetry applications due to the reduced multipath and an improved bit error rate [19]. On-body antennas should also be designed with simple, flexible, miniaturized and reduced radiation characteristics. The gain of these antennae are typically reduced due to the proximity to human tissues [20]. Multiple antenna designs in the literature have also been proposed for WBAN applications [21-26].

The proposed implant antenna has a miniaturized size making it suitable for placement into an arm of the human body model. A three layer tissue model (skin, fat and muscle) which was also used in [9] is used in this research study. The radiating patch of the implant antenna is characterized using a combination of Hilbert fractal and serpentine geometries to realize dual-band operation and satisfactory transmission characteristics inside the arm model. Self-similarity and space filling are the two unique characteristics of Hilbert and serpentine geometries which allow for longer electrical lengths while maintaining the reduced physical size of the radiating patch. These characteristics allow for miniaturization and enhance resonances at the desired frequencies. Kaka et al. (2012) characterized the geometric generation and configuration of the Hilbert fractal geometry-based implant antenna [27]. Zirconia is used as a superstrate material in order to achieve biocompatible properties while preventing direct contact between the implant and human tissue. Along with the proposed implant antenna, on-body antennas for MICS and ISM bands are designed using hexagonal and circular serpentine geometries accordingly.

The first objective is to simulate the proposed implant antenna inside the three-layer human arm model phantom for dual band operation. As a second design objective, the two on-body antennas are characterized and the path loss is analysed. The design methodology of the implant and on-body antennas is introduced in Section 2. Antenna geometry parameterizations are presented in Section 3. Various simulated performance analysis, including return loss, axial ratio and radiation characteristics, are discussed in Section 4. Conclusions are given in Section 5.

II. DESIGN METHODOLOGY OF IMPLANT AND ON-BODY ANTENNAS

The four-step design methodology applied in this research study is shown in Fig. 1. First, the implant antenna was characterized and placed inside a human arm model with dimensions 150 x 75 x 55 mm^3 (width × height × thickness mm^3) for the purpose of determining transmission characteristics, in particular reflection coefficient (S_{11}) values. Modifications are then made on the Hilbert and Serpentine geometries to achieve dual band operation with satisfactory transmission characteristics. S_{11} results better than -10 dB are acceptable for transmission.

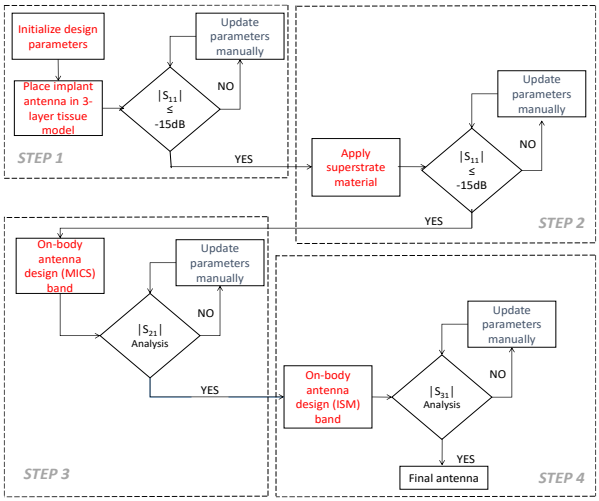


Fig. 1. Proposed generic four–step methodology for implantable antenna design.

The second objective (Step 2) was to consider a superstrate dielectric layer to achieve biocompatible characteristics while preventing direct contact between the implant and human tissues. Additionally, a superstrate can prevent short circuit of the antenna operation by tissue. Some well-known materials that could be used for this approach include Teflon (permittivity – ε_r=2.1; dielectric loss tangent – tan δ=0.001), MACOR

(permittivity – ε_r=6.1; dielectric loss tangent – tan δ=0.005), and ceramic alumina (permittivity – ε_r=9.4; dielectric loss tangent – tan δ=0.006) [1]. The primary challenge here is to select a low-loss, biocompatible material that can insulate the antenna implant with as a thin coating. For this approach, zirconia (permittivity – ε_r=29; dielectric loss tangent – tan δ=0.001) was considered as materials with higher permittivity also resonate at lower frequencies due to shorter effective wavelength [9]. An important research objective in the development of WBAN is to understand the propagation channel characteristics as an essential requirement for the efficient design of wireless communication systems. The third (Step 3) and fourth objectives (Step 4) were to consider the design of on-body antenna designs for both the MICS and ISM bands so that a communication implant with the proposed implant antenna may be realized, as can be seen in Fig. 2. Anatomical tissue models, with the electrical properties of human body tissues, are proposed in the literature [1-3, 5-9, 11-14 and 28]. The proposed antenna in reference [28] is assumed to be implanted subcutaneously into the human model between the shoulder and the elbow with model dimensions of 180 mm x 60 mm x 60 mm. Grey layer, brown layer, green layer and blue layer represents muscle, fat, skin and air layers accordingly. In order to obtain more realistic results in the research study, a four-layer human arm model was used (air, skin, fat and muscle) with the dimensions of 150 mm x 75mm x 55 mm (width × height × thickness mm^3).

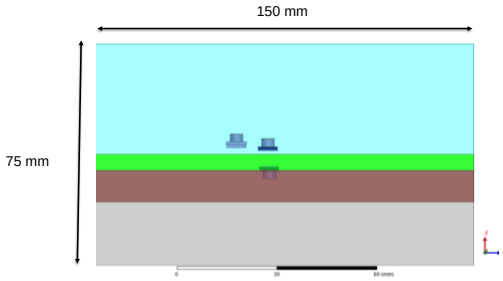


Fig. 2. Propagation channel characteristics for dual band operation from implant antenna to on-body antennas. Grey, brown, green and blue represent muscle, fat, skin, and air, respectively.

III. PROPOSED ANTENNA GEOMETRY PARAMETERIZATION

A. Implant antenna parameterization

Three layers were used to design the proposed antenna: layer 1 was used as a ground plane and layer 2 was characterized as a radiating element. A superstrate material was applied in layer 3 realization of biocompatibility. The final implant antenna geometry, a 3D render of the proposed antenna design, and the three-layer human tissue model that was used to simulate antenna performance, are depicted in Figs. 3 (a), (b) and (c), respectively. Rogers R03010 was used as the substrate with a dielectric constant (ε_r) of 10.2 and loss tangent (tan δ) of 0.0035. The thickness of the substrate was 0.9 mm and the feed position is located at the centre (*xy*-axis) of the structure as depicted in Fig. 3 (b).

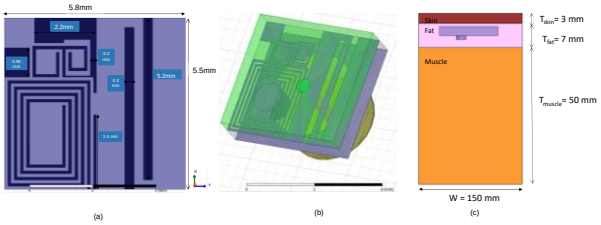


Fig. 3. (a) Final antenna geometry, (b) 3D of proposed antenna design, and (c) three-layer human tissue model.

As can be seen from Figs. 3 (a) and (b), two major design guidelines, Hilbert fractal geometry and serpentine geometry, were applied to achieve the proposed antenna's radiating element geometry. The initial shape of a square was used due to its ease of modification and the fact it is centro-symmetric on (0, 0). The initial square was divided into two vertical sections on the x-axis and a Hilbert fractal geometry modification was applied to the top left of the half-square region. The Hilbert generation process has previously been considered in [27] and the original first iteration modified by removing the right top part of the geometry. The serpentine geometry was applied beneath the Hilbert fractal geometry in order to increase the capacitive nature of the design [9]. The modification for the left-most region was intended to achieve resonance for MICS bandwidth. By utilizing modifications on the left-part of the square, a longer electrical length was achieved resulting in resonance at lower frequencies. Finally, two slits were removed from the centre of the right-part of the square in order to achieve resonance for the ISM bandwidth.

A shorting pin with a radius of 0.5 mm was used at the centre of the serpentine geometry, as seen in the left-part of the design in Fig. 3 (b), to connect radiating elements with the ground plane. This application causes an increase in the effective size of the antenna while also enhancing resonances on the desired frequencies [6]. A coaxial probe is employed as a feeding line and proposed implant antenna parameters were characterized to achieve 50 ohm impedance matching. Figure 3 (b) depicts the Zirconia superstrate material in green (thickness of 0.5 mm) which was used to prevent direct the contact of human tissue by the implant. A three layer tissue model (skin, fat and muscle) was used as a simulation environment with similar electrical characteristics as human arm model. As depicted in Fig. 3 (c), this model has dimensions 150 × 75 × 60 mm^3 (width × height × thickness). The electrical properties of each layer are

specified in Table 1 at 433 MHz and 2.45 GHz. It should be noted that the values are defined by the Office of Engineering and Technology (OET Bulletin 65 Supplement C) of the FCC (2001) [29].

As fat tissue exhibits lower loss characteristics than skin and muscle, the proposed implant was placed into a fat layer at a depth of 5 mm from the surface of the skin [9]. As Zirconia has a higher conductivity than human tissue, the MICS bandwidth resonance is enhanced.

Table 1: Electrical properties of body tissue at the different frequency bands [29]

	Relative Permittivity ε_r		Conductivity σ (S/m)	
Frequency (MHz)	433	2450	433	2450
Body equivalent tissue	56.7	52.7	0.94	1.95
Skin	46.08	38.01	0.7	1.46
Fat	5.57	5.28	0.04	0.1
Muscle	56.87	52.73	0.8	1.74

B. MICS band on-body antenna parameterization

The proposed on-body antenna for operation at the MICS bandwidth was realized using two design concepts. The antenna was designed on a rectangular Roger RO3010 substrate with a dielectric constant (εr) of 10.2 and loss tangent (tan δ) of 0.0035. First, a serpentine geometry was applied on a hexagonal geometry and Zirconia material was applied as a superstrate to shield all sides of the patch and achieve resonance in the desired MICS bandwidth which is necessary due to the higher permittivity of human tissue. Probe-feed excitation device was placed at the centre of the radiator. The initial design geometry is illustrated in Fig. 4 (a) and dimensions are presented in Table 2. The substrate thickness is 0.8 mm and the total antenna size is 6 × 5.7 x 1.3 mm^3 (width × height x thickness). The top layer superstrate thickness is 0.5 mm while the two sides are 0.1 mm each and dark blue area superstrate thickness applied is 5.6 mm.

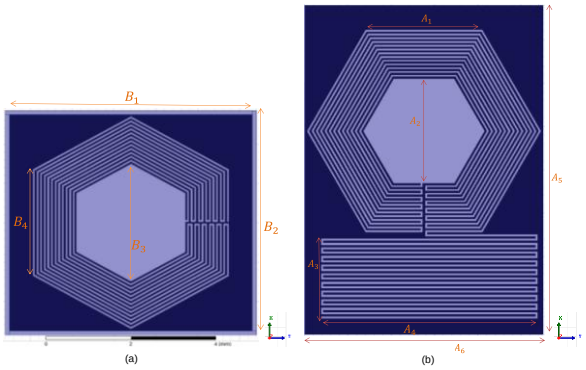


Fig. 4. (a) Initial design geometry of the MICS band on-body antenna, and (b) final design geometry with further increased electrical length.

Table 2: MICS band on-body antenna initial design parameters

Parameters (mm)	B_1	B_2	B_3	B_4
	6	5.7	3	2.7

As a result of the shielding effect and the high electrical permittivity of the Zirconia superstrate, it was observed that the antenna experiences poor radiation efficiency and increased path-loss. The superstrate was removed from the final design geometry in an attempt to rectify these issues and improve the antenna transmission properties, reduce path-loss characteristics and simplify the structure. In addition to this, electrical path length was increased by applying additional rectangular serpentine geometry beneath the hexagonal shape, as shown in Fig. 4 (b). This resulted in lower resonating frequencies as in line with the outcome from first design concept. The final design antenna parameterization is shown in Table 3 with the dimension 5.5 x 7.6 x 0.8 mm^3 (width × height × thickness).

Table 3: MICS band on-body antenna final design parameters

Parameters (mm)	A_1	A_2	A_3	A_4	A_5	A_6
	2.7	2.725	1.95	5	7.6	5.5

C. ISM-band on-body antenna parameterization

The second proposed on-body antenna that operates at the ISM bandwidth was achieved by applying concentric circles slots within a square-form radiator as shown in Fig. 5. The mathematical expression for circular concentricity is shown in equation (1) where *(h,k)* represents the coordinates of the antenna geometry's centre, *r* is the radius of the most inner circle and r_1 is the radius of the second circle. The concentric design was deployed to increase the antenna's electrical length:

$$(x-h)^2+(y-k)^2 = r^2 \; is \; (x-h)^2+(y-k)^2 = r_1^2, (r_1 \neq r). \quad (1)$$

The antenna was again designed on a rectangular Roger RO3010 substrate. The probe-feed excitation device is placed at the centre of the radiator. The central position of the probe is intended to ensure that the probe size has no impact on the spatial dimension of the patch antenna.

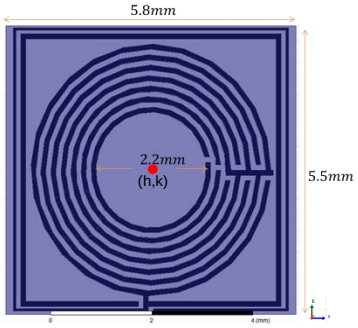


Fig. 5. Design geometry of the ISM band on-body antenna.

IV. SIMULATION RESULTS

Numerical Finite Element Methods (FEM) were applied for the analysis of transmission characteristics of the implant and on-body antennas using commercially available full-wave 3D Ansoft HFSS software. Boundary conditions and meshing of the proposed antennas were chosen with respect to the resonant frequencies with perceived boundary wavelengths. In order to achieve lower resonant frequencies, the use of a dielectric substrate with high permittivity is required in order to achieve shorter effective wavelengths. Electric field distribution for the dual band operation implant antenna is shown in Fig. 6.

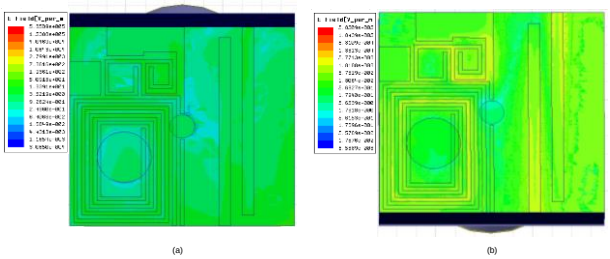


Fig. 6. Implant antenna electric field distribution (a) MICS band and (b) ISM band.

It could be seen from Fig. 6 that the modified Hilbert and serpentine radiator (left hand side of the antenna) was primarily used to achieve a frequency resonance in the MICS band while the two slits on the right-hand side are for the ISM band as expected. There are, however, coupling effects between both radiators. With the help of this dual-band operation, the antenna may be controlled to enter sleep or wake-up modes. The antenna is expected to communicate on MICS bands as soon as a wake-up signal is received in the ISM band. Such an implant would then have improved energy efficiency and a longer life cycle. Figures 7 (a) and (b) display the implant antenna current distribution at MICS and ISM bands respectively showing the same achievement for antenna resonances.

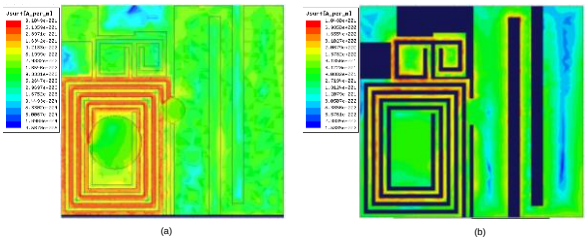
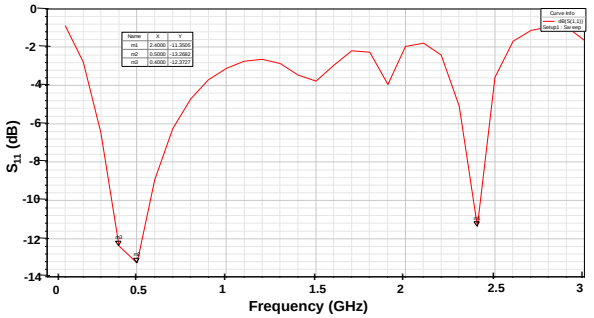
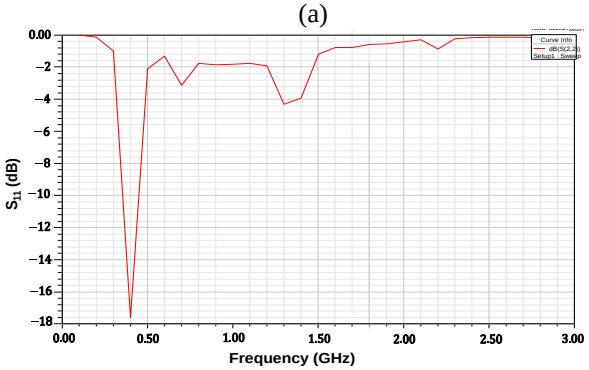


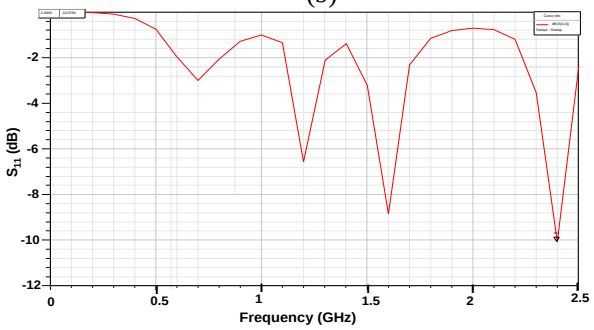
Fig. 7. Implant antenna current distribution (a) MICS band and (b) ISM band.



(a)



(b)



(c)

Fig. 8. Monitoring S_{11} characteristics of (a) proposed antenna with dual band operation, (b) MICS (402-405 MHz) band on-body antenna, and (c) ISM (2.4-2.48 GHz) band on-body antenna.

Figure 8 depicts the simulation analysis of the proposed implant and on-body antenna performances by monitoring simulated reflection coefficient (S_{11}) variations over the transmission bandwidths. It should be noted that the simulation of reflection coefficient analysis of the implant antenna over the MICS (402-405 MHz) and ISM (2.4–2.48 GHz) bands are achieved at the same time and that reflection coefficient characteristics are enhanced with S_{11} below -10 dB as shown in Fig. 8 (a). In addition, the analysis of the MICS and ISM band on-body antennas reveal satisfactory reflection coefficient characteristics as shown in Figs. 8 (b) and (c)

respectively. Voltage Standing Wave Ratio (VSWR) simulation results of the proposed implant antenna and on-body antennas, over transmission bandwidths, are less than 2. These results indicate satisfactory matching characteristics and correlate with the reflection coefficient measurements.

The real and imaginary parts of the input impedance as a function of proposed implant antenna frequency, MICS band on-body antenna and ISM band on-body antenna are shown in Figs. 9 (a), (b) and (c) respectively.

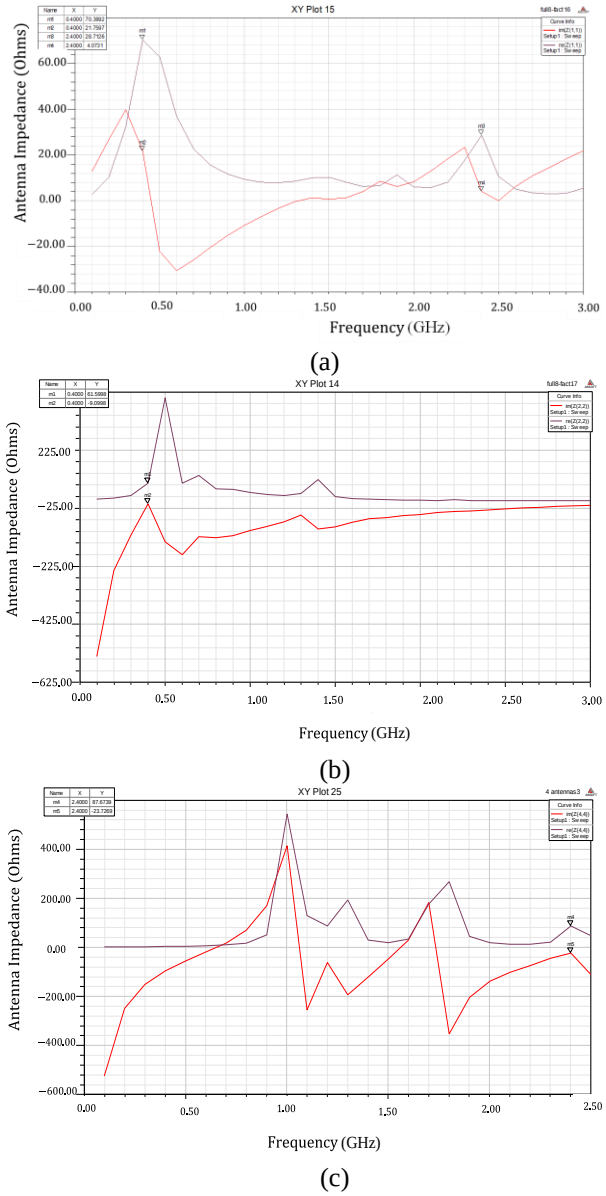


Fig. 9. Input impedance of (a) proposed implant antenna, (b) MICS band on-body antenna, and (c) ISM band on-body antenna: real part (purple line) and imaginary part (red line) over MICS and ISM bands.

It can be observed from Fig. 9 (a) that at the resonance frequency of 400 MHz, the real part of the input impedance is approximately 70 Ω, which is in line with a reflection coefficient of better than -10 dB and an acceptable level for communication. However, the imaginary part is far from zero at 21 Ω. For the ISM band, the real and imaginary parts are approximately 28 Ω and 4 Ω, respectively. On-body antennas can be matched to an approximately 50 Ω line for the mono-band and dual band operations bandwidth. A matching circuit can be used in practice to realize stable 50 Ω matching.

Figures 10 (a) and (b) illustrates the 3D radiation patterns for the 400 MHz and 2.4 GHz resonance frequencies, respectively. Omni-directional radiation properties, with a toroidal shape, were seen in both H-plane and in the E-plane for dual-band operation. Hence, accurate orientation alignment is demonstrated.

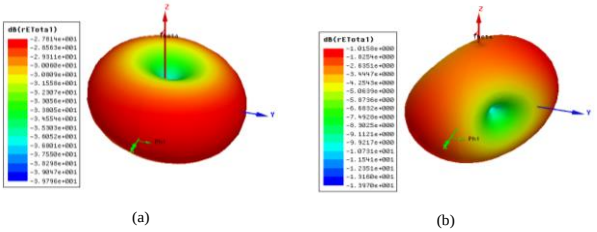


Fig. 10. 3D simulated radiation patterns of the proposed implant antenna at (a) 400 MHz and (b) 2.4 GHz.

Specific Absorption Rate (SAR) is a vital measure of rate and which must adhere to the FCC approved limit [9]. Based on the HFSS simulations, a 10g average SAR from the arm model is given at 105W/kg at 400 MHz and 464 W/kg at 2.4 GHz. It should be noted that SAR values were achieved under a 1W input power condition, illustrated in Fig. 11 for dual band operation. In accordance, the incident power for the proposed implant antenna should be lower than 15.2 mW and 3.4 mW for the MICS and ISM bands, respectively, in order to meet the 1.6 W/kg regulation limit.

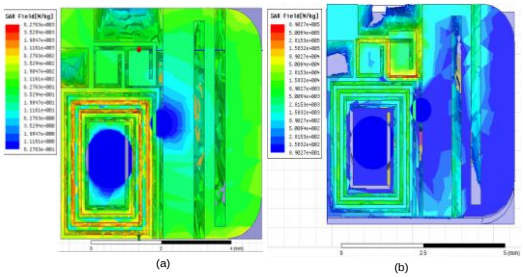


Fig. 11. Proposed implant antenna showing the electromagnetic energy (SAR profile) distribution at (a) 400 MHz and (b) 2.4 GHz.

In addition, SAR characteristics of the on-body antennas (MICS and ISM) are presented in Fig. 12. The same power ratio to energy absorption is applicable for the on-body antennas. Based on this, the peak incident power should not exceed 0.25mW and 0.2mW for the ISM band and MISC band antennas, respectively.

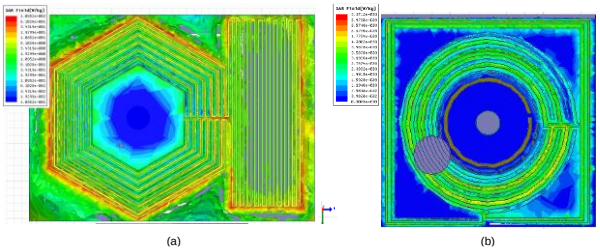


Fig. 12. Proposed on-body antenna showing the electromagnetic energy (SAR profile) distribution at (a) 400 MHz and (b) 2.4 GHz.

In addition to the dual band operation of the implant antenna, the analysis of path loss and propagation channel characteristics of WBAN is also a key research objective as the radio waves must travel from the implant antenna to on-body antennas through human tissue. S_{12} performance, the reverse transmission coefficient between the two antennas used to calculate coupling between antennas, was simulated for the MICS band on-body antenna design at multiple distances above the three-layer human tissue model for diversity reception with results presented in Fig. 13. The frequency behaviour of the S_{12} characteristics is considered to be as important as the S_{11} characteristics as this represents the wireless channel frequency response.

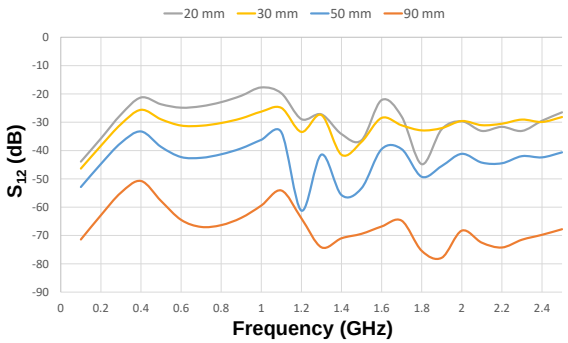


Fig. 13. S_{12} frequency response between the MICS band on-body antenna design and implant antenna at different distances 'd' between antennas.

It is well-known that large distances impose high path loss between implant and on-body antennas. As can be seen in Fig. 13, S_{12} for the MICS band on-body antenna varies from -20 dB to -50 dB as the on-body antenna moves from 90 mm to 2 mm between the three-layer tissue model.

S_{13} performance was also simulated for the ISM band on-body antenna design as shown in Fig. 14. It should be noted again that the ISM band on-body antenna was considered for the purpose of wake up the implant antenna for wireless telemetry transmission.

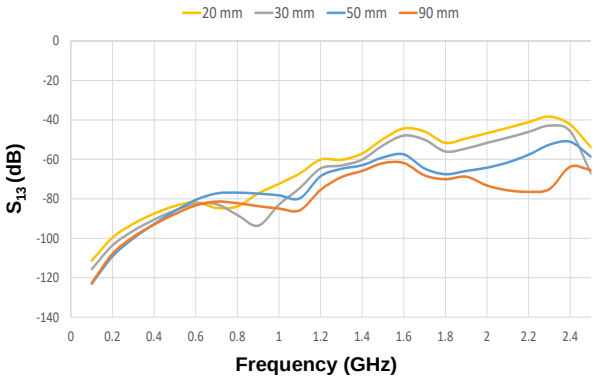


Fig. 14. Monitoring S_{13} characteristics between ISM band on-body antenna design and implant antenna at different distances *d* between antennas.

Figure 14 illustrates the coupling relationship between the in-body dual-band antenna and the ISM on-body antenna. The coupling between the two antennas at the ISM band is dampened compared to the MICS band shown in Fig. 13. This is due to the higher Q-factor experienced by patch antennas at higher frequencies resulting from increased surface wave excitation.

Analysis of the radiation efficiency and gain of the implant antenna suggests that both the gain and efficiency suffer largely as a result of being surrounded by the high-loss human tissue model. A comparison between the proposed implant antenna and other related published studies in terms of physical size, transmission bandwidth and peak gain (dB) is presented in Table 4.

Table 4: Comparison between implant antenna and related studies

Reference	Dimensions (mm x mm x mm)	Bandwidth ($\|S_{11}\|$ < -10 dB)	Peak Gain (dBi)
[11]	22.5 x 22.5 x 2.5 (1265.6 mm^3)	MICS	-2.4
		ISM	-7.5
[30]	16.5 x 16.5 X 2.54 (691.515 mm^3)	MICS	-31
		ISM	-9
[31]	13.4 x 16 x 0.835 (179 mm^3)	MICS	-30.6
		ISM	-19.1
[32]	10.02 x 10.02 x 0.675 (67.8 mm^3)	MICS	-30.5
		ISM	-19.2
[33]	8.75 x7.2 x 0.5 (31.5 mm^3)	MICS	-39.1
		ISM	-21.2
Proposed dual band design	6 x 5.7 x 1.3 (44.6 mm^3)	ISM	-22.7
		MICS	-39.8

The data presented in Table 4 shows that the proposed dual band implant antenna has the smallest physical dimensions and lower peak gain characteristics when compared to related studies. Low peak gain is observed due to the fact that the size reduction results as a less efficient radiator as compared to the λ/2 requirement. Generally, a reduction in antenna efficiency has a direct impact on antenna gain. Importantly, properties including phantom size, electric properties of the tissue model and the placement of implant antenna in tissue model may be differ between publications. Peak gain for the MICS band and ISM band on-body antennas were recorded as -35.6 dBi and -23 dBi, respectively. Low on-body antenna gains were anticipated due to the miniature size of the antennas. The design of low power implant antennas is crucial for patient safety and to eliminate interference with other systems. A maximum of -16 dBm effective radiated power may be used for implant devices in the MICS band in order to not interfere with the collocated Meteorological Aids Service band [33].

V. CONCLUSION

A miniaturized implant and two on-body antennas were proposed in this research study for wireless health monitoring with dual band operation. Hilbert fractal and serpentine radiator geometries were modified and characterized in order to benefit from space filling and self-similar properties. The implant antenna, and on-body MISC (402-405 MHz) and ISM (2.4-2.48 GHz) band antennas have dimensions 5.5 x 7.6 x 0.8 mm^3, 6 x 5.7 x 1.3 mm^3 and 5.8 x 5.5 x 0.8 mm^3 (width x height x thickness mm^3). The ISM band could be used to transmit the wake-up signal for the implant antenna while health monitoring could be completed over the MICS band. A proposed dual-band operation model was simulated in a 150 x 75 x 55 mm^3 four-layer tissue model (air, skin, fat and muscle) with satisfactory transmission characteristics (S_{11} below -15 dB). Zirconia material was used to achieve biocompatible characteristics while preventing a short-circuit between the antenna and human tissue. A coaxial probe will be used as a feeding line for future experimental analysis and a shorting pin was used to connect the radiating elements to the ground plane in order to achieve resonance at specific frequencies. Analysis demonstrated that the implant and on-body antennas have satisfactory, omni-directional S_{12} and S_{13} transmission characteristics and acceptable SAR results were achieved for dual-band operation. The achievement of stable transmission properties with antennas that are notably smaller than those presented in previous studies is a significant achievement of this work.

REFERENCES

[1] A. Kiourti and K. S. Nikita, "A review of implantable patch antennas for biomedical telemetry: Challenges and solutions," *IEEE Trans. Antennas Propag*., vol. 54, pp. 210-228, 2012.

[2] A. Ahmed, T. Kalsoom, M. Rehman, N. Ramzan, S. Karim, and Q. Abbasi, "Design and study of a small implantable antenna design for blood glucose monitoring," *Appl. Comput. Electrom.*, vol. 33, no. 10, pp. 1146-1151, 2018.

[3] Y. Hao, A. Alomainy, P. S. Hall, Y. I. Nechayev, C. G. Parini, and C. C Constantinou, "Antennas and propagation for body-centric wireless communications," *IEEE/ACES International Conference on Wireless Communications and Applied Computational Electromagnetics*, 2005. 10.1109/WCACEM.2005.1469656

[4] C. M. Furse, "Biomedical telemetry: Today's opportunities and challenges," *IEEE Workshop on Antenna Technology Small Antennas and Novel Metamaterials*, Santa Monica, CA, 2009. 10.1109/ IWAT.2009.4906963.

[5] J. Kim and Y. Rahmat-Samii, "Implanted antennas inside a human body: Simulations, designs, and characterizations," *IEEE Trans. Microw. Theory Tech.*, vol. 52, pp. 1934–1943, 2004.

[6] P. Soontornpipit, C. M. Furse, and Y. C. Chung, "Design of implantable micro strip antennas for communication with medical implants," *IEEE Trans. Microw. Theory Tech*., vol. 52, pp. 1944-1951, 2004.

[7] P. Soontornpipit, C. M. Furse, and Y. C. Chung, "Miniaturized biocompatible micro strip antenna using genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 1939-1945, 2005.

[8] C. M. Lee, T. C. Yo, F. J. Huang, and C. H. Luo, "Dual-resonant Pi-shape with double L-strips PIFA for implantable biotelemetry," *Electron. Lett*., vol. 44, pp. 837-838, 2008.

[9] A. O. Kaka, M. Toycan, and S. D. Walker, "Miniaturized stacked implant antenna design at ISM band with biocompatible characteristics," *COMPEL*, vol. 34, pp. 1270-1285, 2015.

[10] Federal Communications Commission, "§ 95.628 Med radio transmitters in the 413-419 MHz, 426-432 MHz, 438-444 MHz, and 451-457 MHz and 2360-2400 MHz bands," in "Code of Federal Regulations, Title 47—Telecommunication, Chapter 1—FCC, Subchapter D-Safety and Special Radio Services, Part 95—Personal Radio Services, Subpart E—Technical Regulations," Tech. Rep. 95.628, 2012.

[11] T. Karacolak, A. Z. Hood, and E. Topsakal, "Design of a dual-band implantable antenna and development of skin mimicking gels for continuous glucose monitoring," *IEEE Trans. Microw. Theory Tech.*, vol. 56, pp. 1001-1008, 2008.

[12] C. J. Sańchez-Fernańdez, O. Quevedo-Teruel, J. Requena-Carrión, L. Inclań-Sańchez, and E. Rajo-

Iglesias, "Dual-band micro strip patch antenna based on short-circuited ring and spiral resonators for implantable medical devices," *IET Microw. Antenna P.*, vol. 4, pp. 1048-1055, 2010.
[13] T. Karacolak, R. Coope, J. Butler, S. Fisher, and E. Topsakal, "In vivo verification of implantable antennas using rats as model animals," *IEEE Antennas Wirel. Propag. Lett.*, vol. 9, pp. 334-337, 2010.
[14] Z. G. Liu and Y. X. Guo, "Dual band low profile antenna for body centric communications," *IEEE Trans. Antennas Propag.*, vol. 61, pp. 2282-2285, 2013.
[15] B. Sanz-Izquierdo, J. A. Miller, J. C. Batchelor, and M. I. Sobhy, "Dualband wearable metallic button antennas and transmission in body area networks," *IET Microw Antenna P.*, vol. 4, pp. 182-190, 2010.
[16] P. J. Soh, G. A. E. Vandenbosch, S. L. Ooi, and M. R. N. Husna, "Wearable dual-band Sierpinski fractal PIFA using conductive fabric," *Electron. Lett.*, vol. 47, pp. 365-367, 2011.
[17] P. J. Soh, G. A. E. Vandenbosch, S. L. Ooi, and N. M. A. Rais, "Design of a broadband all-textile slotted PIFA," *IEEE Trans. Antennas Propag.*, vol. 60, pp. 379-384, 2012.
[18] S. Agneessens and H. Rogier, "Compact half diamond dual-band textile HMSIW on-body antenna," *IEEE Trans. Antennas Propag.*, vol. 62, pp. 2374-2381, 2014.
[19] B. Y. Toh, R. Cahill, and V. F. Fusco, "Understanding and measuring circular polarization," *IEEE T. Educ.*, vol. 46, pp. 313-318, 2013.
[20] C. P. Deng, X. Y. Liu, Z. K. Zhang, and M. M. Tentzeris, "A miniascape-like triple-band monopole antenna for WBAN applications," *IEEE Antennas Wirel. Propag. Lett.*, vol. 11, pp. 1330-1333, 2012.
[21] G. K. Mkongwa, L. Qingling, and Z. Chaozhu, "Design of the Tri-band UWB microstrip patch antenna for WBAN applications," *Appl. Comput. Electrom.*, vol. 38, pp. 1305-1311, 2019.
[22] S. Zhu and R. Langley, "Dual-band wearable antenna over EBG substrate," *Electron. Lett.*, vol. 43, pp. 141-143, 2007.
[23] S. Zhu and R. Langley, "Dual-band wearable textile antenna on an EBG substrate," *IEEE Trans. Antennas Propag.*, vol. 57, pp. 926-935, 2009.
[24] H. R. Raad, A. I. Abbosh, H. M. Al-Rizzo, and D. G. Rucker, "Flexible and compact AMC based antenna for telemedicine applications," *IEEE Trans. Antennas Propag.*, vol. 61, pp. 524-531, 2013.
[25] N. Chahat, M. Zhadobov, R. Sauleau, and K. Mahdjoubi, "Improvement of the on-body performance of a dual-band textile antenna using an EBG structure," *2010 Loughborough Antennas & Propagation Conference*, pp. 465-468, 2010. 10.1109/LAPC.2010.5666201.
[26] M. Mantash, A. C. Tarot, S. Collardey, and K. Mahdjoubi, "Dual-band CPW-fed G-antenna using an EBG structure," *2010 Loughborough Antennas & Propagation Conference*, pp. 453-456, 2010. 10.1109/LAPC.2010.5666211.
[27] A. O. Kaka, M. Toycan, V. Bashiry, and S. D. Walker, "Modified Hilbert fractal geometry, multi-service, miniaturized patch antenna for UWB wireless communication," *COMPEL*, vol. 6, pp. 1835-1849, 2012.
[28] W. Xia, K. Saito, M. Takahashi, and K. Ito, "Performances of an implanted cavity slot antenna embedded in the human arm," *IEEE Trans. Antennas Propag.*, vol. 57, pp. 894-898, 2009.
[29] Federal Communications Commission Office of Engineering and Technology (2001) Evaluating compliance with FCC guidelines for human exposure to radio frequency electromagnetic fields. upplement C: Additional information for evaluating compliance of mobile and portable devices with FCC limits for human exposure to radio frequency emissions.
[30] C. Liu, S. Xiao, Y. X. Guo, Y. Y. Bai, M. C. Tang, and B. Z. Wang, "Compact circularly-polarized microstrip antenna with symmetric-slit," *Electron. Lett.*, vol. 48, pp. 195-196, 2012.
[31] Z. Duan, Y. X. Guo, M. Je, and D. L. Kwong, "Design and in vitro test of a differentially fed dual-band implantable antenna operating at MICS and ISM bands," *IEEE Trans. Antennas Propag.*, vol. 62, pp. 2430-2439, 2014.
[31] L. J. Xu, Y. X. Guo, and W. Wu, "Miniaturized dual-band antenna for implantable wireless communications," *IEEE Antennas Wirel. Propag. Lett.*, vol. 13, pp. 1160-1163, 2014.
[32] Y. Cho and H. Yoo, "Miniaturized dual-band implantable antenna for wireless biotelemetry," *Electron. Lett.*, vol. 52, pp. 1005-1007, 2016.
[33] "International Telecommunications Union-Radio communications (ITU-R), Radio Regulations, SA.1346," ITU, Geneva, Switzerland; available online: http://itu.int/home

Ademola O. Kaka received his OND and HND degrees from the Department of Electrical and Electronic Engineering of Yaba College of Technology, Nigeria in 2008. He has received his M.Sc. and Ph.D. degrees from the Department of Electrical & Electronic Engineering and the department of Computer Engineering of Cyprus International University, North Cyprus in 2012

and 2016 respectively. His research interests include miniaturized, biocompatible and low loss implant antenna designs and their applications in wireless communication systems.

Mehmet Toycan received his B.Sc. degree from the Department of Electric and Electronic Engineering, Eastern Mediterranean University, in 2004. He received the M.Sc. degree in Telecommunications & Information systems and Ph.D. degree in Electronics from University of Essex, Colchester, U.K, in 2005 and 2009 respectively. During his Ph.D. degree, he was a member of Access Networks Laboratory, department of Computer and Electronic Systems at University of Essex. He was also involved in the Network of Excellence, BONE-project ("Building the Future Optical Network in Europe") where he was collaborating with groups from other universities to develop efficient, low-cost optical access network architectures. He is currently working as a Full-time Lecturer and the Director of Communication Technologies Research Centre in Cyprus International University. His research interests are in the fields of cost and energy efficient, scalable optical access networks for FTTH scenarios featuring wireless communication transmission and implant antenna designs for wireless health monitoring applications. In addition, his information systems research team is recently working on electronic health solutions for developing countries.

Stuart D. Walker was born in Dover, UK in 1952. He received the B.Sc. (Hons) degree in Physics from Manchester University, Manchester, UK, in 1973, the M.Sc. degree in Telecommunications Systems and Ph.D. degree in Electronics from Essex University, Colchester, UK, in 1975 and 1981, respectively. He is currently a Full Professor at Essex University where his main research interest is access networks. He has been involved in many EPSRC and European Union research projects over the years, recent examples being BONE, MUSE-I and II, UROOF, STRONGEST, FIVER and OASE. He has published over 250 peer-reviewed journal and conference papers and holds four patents.

Doğa Kavaz received her B.Sc. degree from the Department of Chemistry, Hacettepe University in 2006. She received the M.Sc. degree in Biochemistry (Chemistry) and Ph.D. degree in Nanotechnology (Chemistry) from Hacettepe University, in 2009 and 2013. During her Ph.D. she went to Harvard University as a Researcher for 1 year with a scholarship from TUBITAK. She took part in a joint project with various universities and developed biomaterials for use in tissue engineering. She is currently working as a Full-time Lecturer and Researcher of Biotechnology Research Centre in Cyprus International University. Her research interests include the development and characterization of biomedical materials to be used in medical purposes.

Design and Simulation of 77GHz Substrate Integrated Waveguide Slot Array Antenna

Bo-Yang Guo [1, 2] and Hong Jiang [1]

[1] College of Communication Engineering
Jilin University, Changchun, Jilin Province 130012, China
guoby17@mails.jlu.edu.cn, jiangh@jlu.edu.cn

[2] No. 65024 Unit, People's Liberation Army of China
Dalian, Liaoning Province 116041, China

Abstract ─ In this paper, a 77GHz slot array antenna for millimeter-wave automotive radar application is designed and simulated using HFSS, in which the substrate integrated waveguide (SIW) is adopted. First, a linear array with 16 slots is designed, and its parameters are achieved by extracting the characteristics of the radiating slot in the array environment and comparing them with the array synthesis result. Then, a tree-shaped unequal feeding network with improved vertical transition is presented to make the entire array more compact. Finally, the configuration of the designed 8 × 16 planar array antenna is given, and the fabrication error is analyzed. The simulation results of the 77GHz SIW slot array antenna show that its sidelobe levels (SLLs) are below -25dB in the E-plane and below -27dB in the H-plane, and the corresponding half-power beamwidths are 13.9° and 9.7°, respectively.

Index Terms ─ Automotive radars, feeding network, slot array antenna, substrate integrated waveguide (SIW).

I. INTRODUCTION

77GHz millimeter-wave radar enables high spatial resolution and high transmit power for a small antenna size. Also, it is robust against environmental influences such as extreme temperatures, bad light or weather conditions [1].

Recently, the 77GHz array antenna for millimeter-wave automotive radar application has drawn more and more attention. In [2] and [3], a 25×10 and a 20×16 microstrip series-fed array antennas for millimeter-wave automotive radar were deigned, respectively. However, the ideal sidelobe levels (SLLs) cannot be achieved by the radiation effect of feeding networks and their antenna sizes are large. In [4], a series-fed patch array antenna was presented, which has reduced the size and has a wide bandwidth, nevertheless, its beamwidth in elevation is large that the radar is easy to be affected by ground clutter in practical application. In general, the 3dB beamwidth in elevation is ±5 degree. In [5], a series-fed patch array antenna for 77GHz automotive radar was designed and measured, which has reduced the radiation of the feeding network, however, the effect on the radiation pattern still exists owing to the half-microstrip structure. In fact, all the above-mentioned microstrip patch antennas are not the appropriate choices as the radiation loss and dielectric loss of microstrip are relatively large in the millimeter-wave band.

Substrate integrated waveguide (SIW) is a new planar guided-wave structure fabricated in a dielectric substrate. Its guided-wave characteristic is close to that of conventional metallic waveguide [6]. It keeps both the advantages of metallic waveguide and planar microstrip structure. Waveguide slot array antennas are widely used in radar and communication systems featuring high gain, high efficiency, low cross-polarization levels and great capability of accurate control of the radiation patterns [7]. The SIW slot antenna inherits the virtues of waveguide slot antenna, and also has the advantages of low profile, compact size, low cost, and ease of being integrated with planar circuits [8]. In [9], an SIW slot array with flat-shoulder-shaped radiation pattern was designed for both medium-range radar and long-range radar applications, in which the desired radiation pattern can be obtained. However, its reliability is to be improved for simultaneous detection of two ranges in one antenna by pattern synthesis.

In this paper, we present a design of a 77GHz SIW slot array antenna for millimeter-wave automotive radar application, which is achieved via EM simulation using HFSS. It owns wide bandwidth, miniaturization, low sidelobes, narrow beams and high gain.

In the classic design of longitudinal slot array, the Elliott's iterative method is usually used, in which all the mutual coupling effects are taken into account, and the slot parameters for the desired aperture distribution and impedance matching are accurately determined with Elliott's iterative procedure [7]. However, the Elliott's

Submitted On: December 17, 2019
Accepted On: April 10, 2020

procedure is computationally complicated. With the development of the electromagnetic (EM) computation techniques, it is possible to design a slot array antenna using EM simulation software like HFSS.

Unlike the conventional Elliott's iterative method, we exploit the array synthesis result to obtain the design parameters in serial-fed array. First, we design a linear array with 16 slots and obtain its parameters by extracting the characteristics of the radiating slot in the array environment and comparing them with the array synthesis result. Then, we present a tree-shaped unequal feeding network with improved vertical transition, which can increase space utilization and reduce the entire antenna size. Finally, we design the 8 × 16 planar array antenna, and analyze the fabrication errors to test the tolerance of the antenna.

II. DESIGN OF LINEAR ARRAY

The design parameters need to meet the requirements of certain amplitude distribution and impedance matching when we design an SIW linear array with slots. Its design procedure is as follows: The first step is to extract the radiating characteristics of the radiation element, which is the relationship of the slot normalized resonant conductance against resonant slot length and slot offset. The next step is to obtain the excitation amplitude distribution from the array synthesis result, and according to the distribution, the normalized conductance of each slot is determined. The final step is to obtain the design parameters of each slot by comparing it with the characteristic curve of slot.

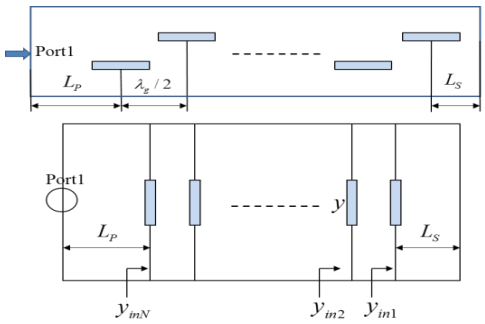


Fig. 1. Slotted waveguide and its equivalent circuit.

According to the principle of the slot characteristics extraction [10], a waveguide with N slots and its equivalent circuit can be shown in Fig. 1. The distance from the first slot to the short-circuit matching stub and the distance from the last slot to the input port are L_s and L_p, respectively.

According to the transmission line theory, for a transmission line of length L_s with an arbitrary load, its normalized input admittance is:

$$y_{in}(L_s)=\frac{y_l+j\tan(\beta L_s)}{1+jy_l\tan(\beta L_s)}, \tag{1}$$

where y_l denotes the normalized admittance of load. When the transmission line has a short-circuit load as in Fig. 1, (1) can be rewritten as:

$$y_{in}(L_s)=\frac{1}{j\tan(\beta L_s)}. \tag{2}$$

The lengths and offsets of all slots are set to be the same, consequently, the admittance of every slot can be assumed with an identical value y. Thus, for the 1st slot, the normalized input admittance y_{in1} can be expressed as:

$$y_{in1}=\frac{1}{j\tan(\beta L_s)}+y. \tag{3}$$

Similarly, y_{in2} and y_{inN} can be derived as:

$$y_{in2}=\frac{y_{in1}+j\tan(\beta\frac{\lambda_g}{2})}{1+jy_{in1}\tan(\beta\frac{\lambda_g}{2})}+y, \tag{4}$$

$$y_{inN}=\frac{y_{inN-1}+j\tan(\beta\frac{\lambda_g}{2})}{1+jy_{inN-1}\tan(\beta\frac{\lambda_g}{2})}+y. \tag{5}$$

As the distance from the Nth slot to the input port is L_p, the input admittance at the input port of waveguide can be expressed as:

$$y_{in}=\frac{y_{inN}+j\tan(\beta L_p)}{1+jy_{inN}\tan(\beta L_p)}. \tag{6}$$

In addition, the normalized input admittance can be also computed from the reflection coefficient [10], i.e.,

$$y_{in}=\frac{1-\Gamma_{11}e^{j2\beta L_p}}{1+\Gamma_{11}e^{j2\beta L_p}}+j\cot(\beta L_s), \tag{7}$$

where Γ_{11} equals to S_{11}, and $\beta=2\pi/\lambda_g$, λ_g is the guided wavelength in equivalent waveguide. Here, when $L_p=2\lambda_g, L_s=\lambda_g/4$, we can have:

$$y_{inN}=y_{in}=\frac{1-S_{11}}{1+S_{11}}. \tag{8}$$

Therefore, the relationship can be established between S_{11} of the input port and total normalized admittance of all slots. The only unknown parameter in (8) is S_{11} of the slotted waveguide, which can be obtained by simulation via HFSS. Then the normalized admittance of every single slot can be calculated by:

$$y=\frac{1-S_{11}}{N(1+S_{11})}, \tag{9}$$

where N is the number of slots in array.

In our design, the slot characteristics extraction procedure is carried out in a 5×16 array model as shown in Fig. 2, in which all the mutual coupling effects in the array environment are taken into account as much as possible.

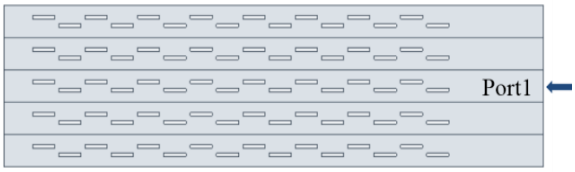


Fig. 2. Equivalent slotted array model for slot characteristics extraction via HFSS.

For parameter extraction, an equivalent dielectric-filled rectangular waveguide model is used according to the SIW equivalent theory to reduce the HFSS meshing and simulation time [11]. The equation of the equivalent width can be expressed as:

$$a_{RWG} = a - 1.08\times\frac{d^2}{p} + 0.1\times\frac{d^2}{a}, \tag{10}$$

where a_{RWG} denotes the width of equivalent rectangular waveguide, and a, d, p denotes the width of the SIW, the period and diameter of the via, respectively.

In the extracting procedure, the slot length is optimized at various offsets. All slots can be considered as in resonance state when the objective function $\mathrm{Im}(S_{11})=0$. By recording the slot length and S_{11} at this point, the resonant lengths and normalized conductance versus various slot offset can be obtained. We can use MATLAB to fit the relationship curve of them. The fitting curves are shown in Fig. 3.

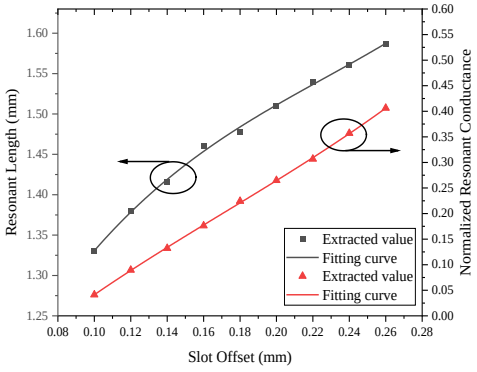


Fig. 3. Characteristic curve of radiating slot.

A certain design margin should be provided for processing in our design and simulation, in order to accommodate unexpected dimension variations in fabrication procedure. A relative strict SLLs standard of -30dB Taylor distribution is adopted for the synthesis in H-plane pattern. The normalized conductance of *n*th slot can be calculated by:

$$g_n = \frac{U_n^2}{\sum_{n=1}^{N} U_n^2}, \tag{11}$$

where U_n is the *n*th element's excitation amplitude from the synthesis. When the normalized conductance of each slot is obtained, the design parameter (length and offset) of each slot can be obtained by comparing it with the slot characteristic curve. The configuration of the SIW slot linear array is shown in Fig. 4.

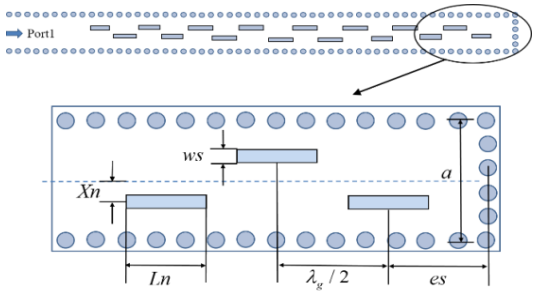


Fig. 4. Configuration of the SIW slot linear array.

In Fig. 4, the width of SIW is 2.3mm, the diameter of the metallized via is 0.4mm, and the distance between two adjacent vias is 0.6mm. The distance between adjacent longitudinal slots is one-half guide wavelength, and the short-circuit matching stub distance is a three-quarter guided wavelength. The liner array is designed on Rogers 5880 dielectric substrate with relative dielectric constant $\varepsilon_r = 2.2$ and thickness 0.508mm. Further optimization is usually required since the simulation results by the initial design parameters are often not perfect. The final optimized design parameters are listed in Table 1.

Table 1: Design parameters of the linear array

Parameter	Value (mm)	Parameter	Value (mm)
L_1	1.3	X_1	0.09
L_2	1.3	X_2	-0.09
L_3	1.31	X_3	0.1
L_4	1.33	X_4	-0.1
L_5	1.36	X_5	0.1
L_6	1.38	X_6	-0.12
L_7	1.4	X_7	0.14
L_8	1.41	X_8	-0.14
es	2.95	*ws*	0.2

The simulation results of S_{11} and radiation patterns of the linear array are shown in Fig. 5 and Fig. 6. The optimized results show that S_{11} is below -10 dB from

76GHz to 78.3GHz, and a relative bandwidth of 2.99% is achieved, with the gain of 15.3 dBi at 77GHz. The half-power beamwidth (HPBW) of the H-plane pattern is 8.9°, and the SLLs are lower than -26 dB. Therefore, the design of SIW serial-fed slot linear array has achieved good performance and will be used as the sub-array of planar array in the next step.

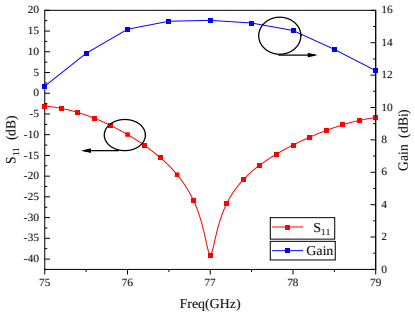


Fig. 5. Simulated S_{11} of the linear array.

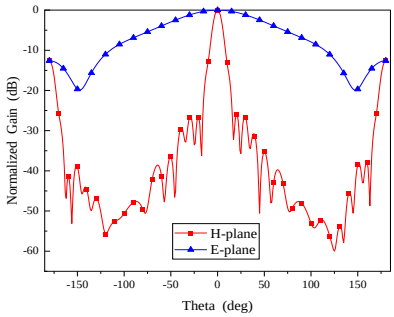


Fig. 6. Simulated radiation patterns of the linear array.

III. DESIGN OF FEEDING NETWORK

Based on the previous design of the linear array, an 8×16 planar array antenna can be designed, in which an eight-way tree-shaped unequal feeding network with transition is required to excite the eight sub-arrays with tapered amplitude and in-phase output.

A. Transition between RWG and SIW

As the transmission loss of coaxial line in millimeter-wave band is relatively large, a transition from rectangular waveguide (RWG) to SIW is generally used as the feed. The commonly-used waveguide-SIW transition is a vertical transition one [12]. A coupling aperture with the size smaller than the waveguide inner wall is etched on the broadside of a slightly wider SIW, which is called the step-SIW, and the standard waveguide is vertically fixed to the coupling aperture. Impedance matching is achieved by adjusting the size of the step-SIW and coupling aperture.

In this section, we first simulate a conventional RWG-SIW transition according to [12], whose configuration and simulated S-parameters results are shown in Fig. 7 and Fig. 8, respectively. The feeding waveguide employed here is a WR-10 waveguide with inner wall sizes 2.54mm and 1.27mm, respectively. The simulation results show that a 12% bandwidth (69.4GHz-78.5GHz) is achieved with $S_{11} < -15dB$ and insert loss lower than 0.3 dB, which can provide smooth transition from RWG to SIW.

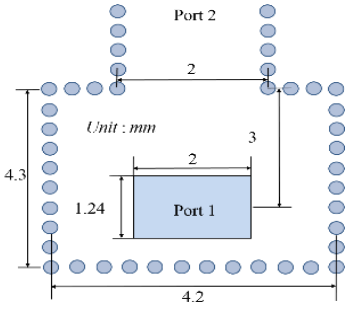


Fig. 7. Configuration of the conventional RWG-SIW transition.

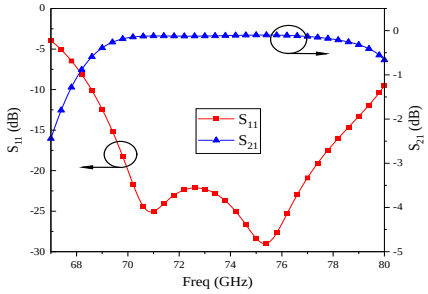


Fig. 8. Simulated S-parameters of the conventional transition.

B. Design of eight-way feeding network with improved vertical transition

The feeding network usually consists of a tree-shaped power divider and a vertical transition. The conventional and proposed feeding networks are shown in the left and right sides of Fig. 9, respectively.

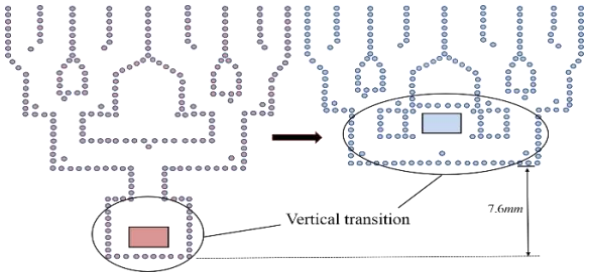


Fig. 9. Configuration of conventional and proposed feeding network.

As shown in Fig. 9, the vertical transition of the conventional feeding network is separately placed below the power divider, and it requires to pass through a certain length's SIW transmission line to achieve stable transmission before connecting to the power divider. Clearly, it makes the feeding network not compact enough. Besides, there is an idle area above the first T-junction, and the space utilization is too low.

We propose a feeding network with improved vertical transition. The configuration of the novel feeding structure is shown in Fig. 10, in which the vertical transition is integrated in the idle area of power divider. The size of the coupling aperture and the width of the step-SIW can refer to the design results of the transition in Section III-A. The upper-wall's distance can be obtained through simple optimization.

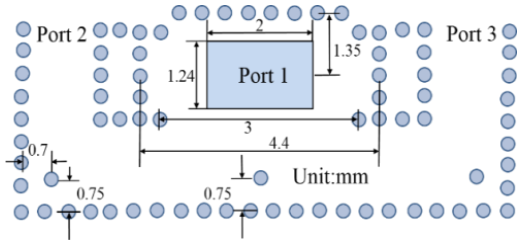


Fig. 10. Configuration of the improved transition.

The HFSS model measurement shows that by this change, the geometry of power divider has reduced from 18.5mm×18.4mm to 10.9mm×18.4mm, which has a 40% reduction. Furthermore, the original idle area is reutilized, thus the space utilization has increased.

The improved transition is next to a T-junction which has an equal power dividing ratio to the port 2 and port 3. The simulation result of S-parameters with this structure is shown in Fig. 11, it can be seen that a 16% bandwidth (70.6GHz-82.6GHz) is achieved with $S_{11} < -15dB$ and insert loss lower than 0.4 dB, which validates that the proposed transition can provide smooth transition to the SIW feeding network.

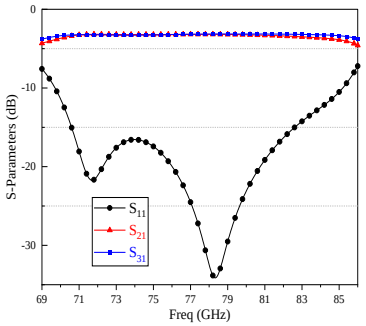


Fig. 11. Simulated S-parameters of the improved transition.

To achieve the low SLLs in the E-plane, eight ways in-phase and tapered amplitude output are required. As with the linear array, a design margin is provided for processing, and -30 dB Taylor distribution is adopted for the synthesis in E-plane pattern. The power divider is needed to divide into three stages, consisting of three T-junctions, four Y-junctions with different power dividing ratios and in-phase output, and six L-junctions at each corner to suppress reflection. These structures need to be designed separately.

The configurations of Y-junction and T-junction with uneven power dividing ratio are shown in Fig. 12. By adjusting the position of the matching via, we can improve the power reflection [13] and control the power dividing ratio, while the phase shift can be slightly tuned by adjusting the width of SIW branch [14], as different widths of SIW have different phase constants [15].

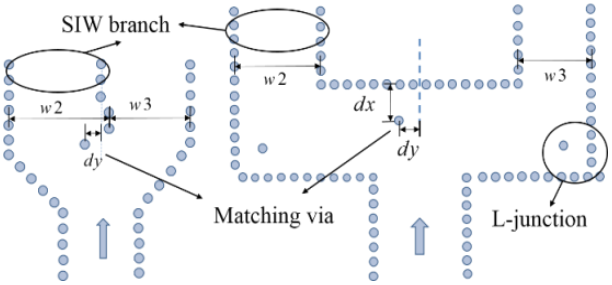


Fig. 12. Configurations of Y-junction and T-junction with a certain dividing power ratio.

We combine the optimized Y-junctions and T-junctions of each stage together to form the power divider and simulate in HFSS. Simulation results of the amplitude and phase characteristics of the complete eight-way feeding network are shown in Fig. 13 and Fig. 14. The results show that the output power dividing ratio of eight ports at 77GHz basically meets the desired goal, and the output phase error is ± 3.5 degree, which can be regarded as in-phase output.

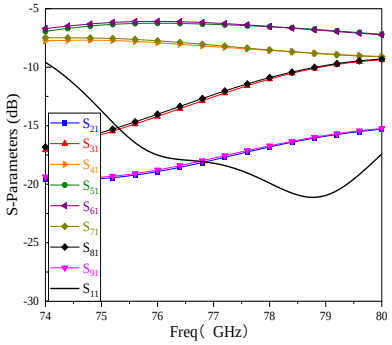


Fig. 13. Simulated amplitude characteristics of the eight-way feeding network.

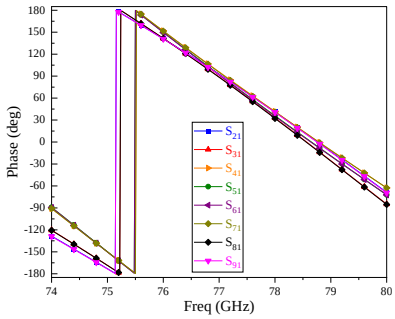


Fig. 14. Simulated phase characteristics of the eight-way feeding network.

IV. DESIGN OF 8×16 PLANAR ARRAY

The configuration of the designed 8×16 SIW slot planar array is shown in Fig. 15. The waveguide feeding port is located on the back of the antenna. The performance simulation results of the antenna are shown in Fig. 16-Fig. 19.

It can be seen from simulation results that S_{11} is below -10 dB from 76.4GHz to 78.2GHz, a relative bandwidth of 2.34% is achieved, with the gain of 23.6 dBi at 77GHz. The E-plane pattern has a half-power beamwidth of 13.9° at 77GHz, the SLLs are lower than -25 dB, and the H-plane's beamwidth is 9.7°, the SLLs are lower than -27 dB. Good performance has been achieved such as low SLLs, narrow beams and high gain.

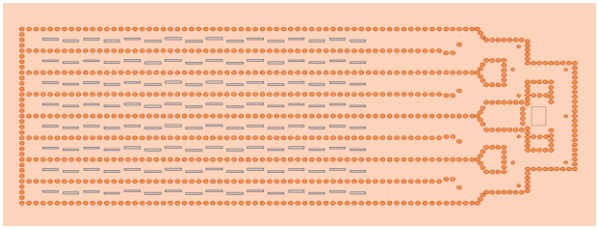


Fig. 15. Configuration of the 8×16 SIW slot array antenna.

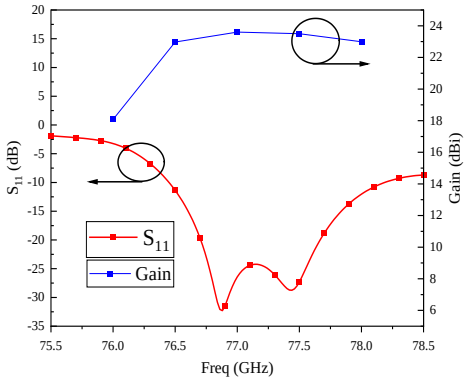


Fig. 16. Simulated S_{11} and gain of the designed antenna.

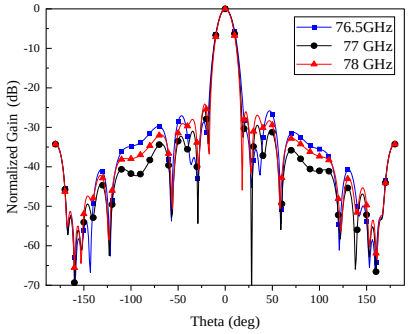


Fig. 17. Simulated E-plane patterns of the designed antenna.

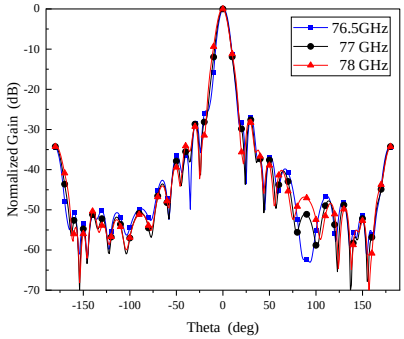


Fig. 18. Simulated H-plane patterns of the designed antenna.

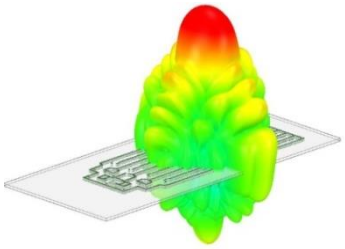


Fig. 19. 3D beam pattern of the designed antenna.

V. ERROR ANALYSIS

In the millimeter-wave band, the performance of device is very sensitive to its dimension. When the antenna is fabricated, a tiny change of design dimension may greatly affect the antenna performance. It is necessary to perform an error analysis to test the tolerance of the antenna.

Considering that the error analysis of the entire antenna will significantly increase the complexity and time consumption, the analysis is carried out in SIW linear array, in which the length error and offset error of radiation slots are taken into account and their impact on performance is analyzed and discussed, respectively.

Generally, the machining error is controlled within ±0.02mm. Therefore, the error analysis of 16 slots in

linear array can be divided into the following four cases [16]: all slots in positive tolerance, i.e., +0.02mm, all slots in negative tolerance, i.e., -0.02mm, half of them in positive tolerance with half in negative tolerance and all in random tolerance.

The impact of length error and offset error on H-plane pattern are shown in Fig. 20 and Fig. 21. We can see that the length or offset error will worsen the pattern slightly, mainly reflected in the elevation of sidelobes of 1-2 dB and the broadening of the beamwidth within 1 degree. In general, the H-plane pattern is still in good performance. The length or offset error will cause the shift of resonance point, as shown in Fig. 22 and Fig. 23. This is because that the lengths and offsets in the design are obtained from the slot characteristic curves extracted in the resonance state, and the change of length or offset will inevitably lead to the change of resonance state.

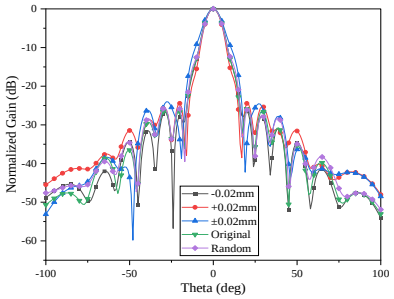


Fig. 20. Impact of length error on H-plane pattern.

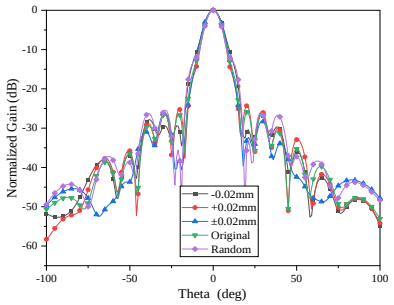


Fig. 21. Impact of offset error on H-plane pattern.

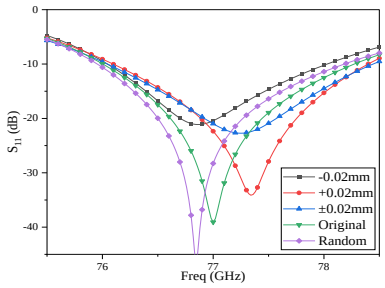


Fig. 22. Impact of length error on S_{11}.

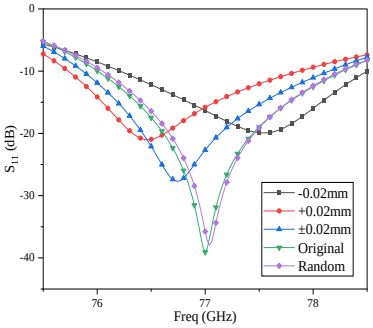


Fig. 23. Impact of offset error on S_{11}.

It can be seen from Fig. 20 to Fig. 23 that the linear array basically keeps good performance in a certain range of errors, so we can infer that the complete antenna should also have a certain tolerance. In addition, it shows that in order to achieve good performance, it is necessary to provide a certain design margin in the design process, and it's better to choose the manufacturing technology with high precision in the future.

VI. CONCLUSION

A 77GHz SIW slot array antenna is designed and simulated using the high-frequency simulation software HFSS in this paper. First, the radiating characteristic parameters of the slot are extracted using an equivalent array model, and a linear slot array with low SLLs in H-plane is designed, then an 8 × 16 planar array fed by a novel tree-shaped feeding network is designed with an improved vertical transition, which can increase the space utilization and reduce the antenna size. The fabrication errors are also analyzed to test the tolerance of the antenna. Simulation results show that the designed antenna array has the characteristics of wide bandwidth, miniaturization, low sidelobes, narrow beams and high gain. It can be served as the transmitting antenna for automotive radar applications.

REFERENCES

[1] J. Hasch, E. Topak, R. Schnabel, et al., "Millimeter-wave technology for automotive radar sensors in the 77GHz frequency band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 845-860, Mar. 2012.

[2] Q. Zhang, L. Wang, and X. Zhang, "Millimeter-wave microstrip comb-Line antenna array for automotive radar," *2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE)*, pp. 1-3, Hangzhou, China, 2018.

[3] C. Yi and W. B. Dou, "Microstrip series fed antenna array for millimeter wave automotive radar applications," *2012 IEEE MTT-S International Microwave Workshop Series on Millimeter Wave*

Wireless Technology and Applications, pp. 1-3, Nanjing, 2012.

[4] X. Shang, F. Zhang, R. He, and H. Zhang, "A 77GHz miniaturized microstrip antenna array for automotive radar," *2019 International Conference on Microwave and Millimeter Wave Technology (ICMMT)*, pp. 1-3, Guangzhou, China, 2019.

[5] J. Xu, W. Hong, H. Zhang, et al., "Design and measurement of array antennas for 77GHz automotive radar application," *2017 10th UK-Europe-China Workshop on Millimetre Waves and Terahertz Technologies (UCMMT)*, Liverpool, UK, Sept. 11-13, 2017.

[6] M. Bozzi, A. Georgiadis, and K. Wu, "Review of substrate-integrated waveguide circuits and antennas," *IET Microwaves, Antennas & Propagation*, vol. 5 , no. 8 , pp. 909-920, June 2011.

[7] J. F. Xu, W. Hong, P. Chen, et al., "Design and implementation of low sidelobe substrate integrated waveguide longitudinal slot array antennas," *IET Microwaves, Antennas & Propagation*, vol. 3 , no. 5, pp. 790-797, Aug. 2009.

[8] S. Moitraa and P. S. Bhowmik, "Effect of various slot parameters in single layer substrate integrated waveguide (SIW) slot array antenna for Ku-Band applications," *The Applied Computational Electromagnetics Society (ACES) Journal*, vol. 30, no.8, pp. 934-939, Aug. 2015.

[9] Y. Yu, W. Hong, H. Zhang, et al., "Optimization and implementation of SIW slot array for both medium- and long-range 77 GHz automotive radar application," *IEEE Transactions on Antennas & Propagation*, vol. 66, no. 7, pp. 3769-3774, July 2018.

[10] M. Q. Qi, W. Wang, M. P. Jin, "A method of calculating admittance of waveguide slot," *2005 Asia-Pacific Microwave Conference Proceedings*, Suzhou, China, Dec. 4-7, 2005.

[11] F. Xu and K. Wu, "Guided-wave and leakage characteristics of substrate integrated waveguide," *IEEE Transactions on Microwave Theory & Techniques*, vol. 53, no.1, pp. 66-73, 2005.

[12] K. D. Wang, W. Hong, and K. Wu, "Broadband transition between substrate integrated waveguide (SIW) and rectangular waveguide for millimeter-wave applications," *Applied Mechanics & Materials*, vol. 130-134, pp. 1990-1993, 2012.

[13] A. Piroutiniya and P Mohammadi, "The substrate integrated waveguide T-junction power divider with arbitrary power dividing ratio," *The Applied Computational Electromagnetics Society (ACES) Journal*, vol. 31, no.4, pp. 428-433, Apr. 2016.

[14] J. Wang and Y. J. Cheng, "W-band hybrid unequal feeding network of waveguide and substrate integrated waveguide for high efficiency and low sidelobe level slot array antenna application," *International Journal of Antennas and Propagation*, vol. 2017, Article ID 7183434, 8 pages, 2017.

[15] Y. J. Cheng, W. Hong, and K. Wu, "Broadband self-compensating phase shifter combining delay line and equal-length unequal-width phaser," *IEEE Transactions on Microwave Theory & Techniques*, vol. 58, no. 1, pp. 203-210, 2010.

[16] R. Liu, "W-band waveguide slot array antenna design," *A Master Thesis from University of Electronic Science and Technology of China*, 2015.

Bo-Yang Guo received the B.S. degree in Communication Engineering from Information Engineering University, China, in 2012. He is currently a graduate student at College of Communication Engineering, Jilin University, China. His current research interest is antenna design and EM simulation.

Hong Jiang received the Ph.D. degree in Communication and Information System from Jilin University, China, in 2005. From 2010 to 2011, she was a Visiting Scholar with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. She is currently a Professor at the College of Communication Engineering, Jilin University, China. She is a member of Institute of Electrical and Electronics Engineers (IEEE) and senior member of the Chinese Institute of Electronics (CIE). Her research fields focus on array signal processing, radar signal processing and antenna design.

A Multiply Parasitic-Coupled, Three-Dimensional Antenna Array with Wide Elevation Angle for Seamless UAV Communications

Dong-Geun Seo and Wang-Sang Lee

Department of Electronic Engineering/Engineering Research Institute (ERI)
Gyeongsang National University (GNU), 501, Jinju-daero, Jinju, Gyeongnam, 52828, Republic of Korea
wsang@gnu.ac.kr

Abstract ─ A multiply parasitic-coupled, three-dimensional antenna array with wide elevation angle for seamless unmanned aerial vehicle (UAV) communications is proposed. The proposed array consists of an upper substrate with a two-element dipole array, multiple parasitic elements, two supports including a microstrip feeding, and a lower substrate with a feeding network and its ground plane as a metallic reflector for wide radiation pattern and reduced back radiation. The proposed array operates with a 18% impedance bandwidth ranging from 4.51 GHz to 5.43 GHz by covering an UAV communications frequency band. Measured peak gain, total efficiency, and half-power beamwidth (HPBW) of the proposed array are approximately 5.5 dBi, 95%, and 140°, respectively.

Index Terms ─ Antenna array, parasitic-coupled, three-dimensional antenna, UAV, wide elevation angle.

I. INTRODUCTION

Since unmanned aerial vehicles (UAVs) are able to remote control without people in aircraft, various applications using UAVs have been researched and expanded and have grown extensively over the years. In the military, UAVs utilize not only surveillance activities but also search-attack missions by focusing on the specific places or target objects. Furthermore, recent radio frequency identification (RFID) technology utilizes the UAVs for indoor or outdoor warehouse inventory management system in civilian industry. To realize these applications, the development of the customized antenna for reliable and stable UAV-based wireless communications is required. First of all, the UAV antenna requires a compact size such as lightweight and low-profile structure for stable long-time flight and communications. Secondly, it needs to have a symmetrical structure to balance the aircraft. Moreover, high gain and wide beam coverage are required to seamless UAV to UAV or UAV to infra communications. To satisfy these demands, previous researches have been studied in [1]–[7]. Crossed dipole loaded with magneto-electric dipole antenna is proposed in [1], [2]. The antenna with a metallic cavity has a wide axial-ratio (AR) beamwidth and a high front-to-back ratio [3]. In [4] and [5], the low-profile antenna with wide beamwidth is proposed utilizing the vertical currents. Using reconfigurable feeding network [6] and shorted vertical plates [7], the wide beams are generated. Although these antennas have achieved wide beamwidth and a low-profile structure, they have a low productivity and are not suitable for use in UAVs because their structures are complicated. In this paper, a lightweight, low-profile dipole antenna array with wide elevation angle using multiple parasitic elements is proposed for seamless UAV communications.

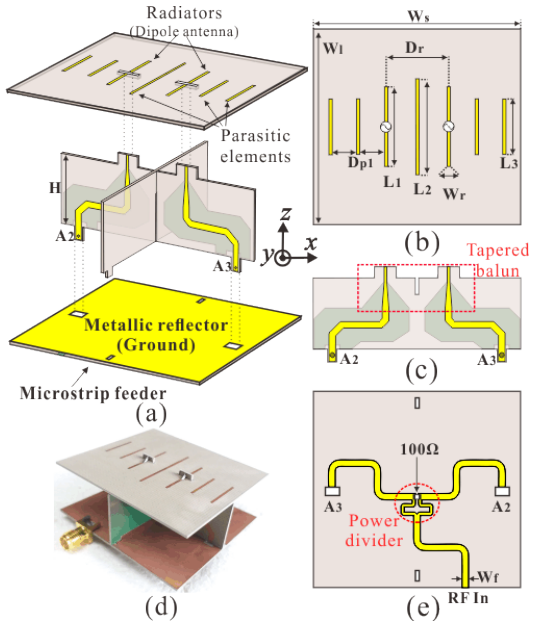


Fig. 1. Configuration of the proposed antenna array and its fabrication: (a) a perspective view of the proposed antenna array, (b) a top view of the upper substrate with dipole array and parasitic elements, (c) a top view of the antenna support with microstrip lines and a tapered balun, (d) the proposed array prototype, and (e) a bottom view of the lower substrate with a microstrip feeder.

Submitted On: December 26, 2019
Accepted On: April 12, 2020

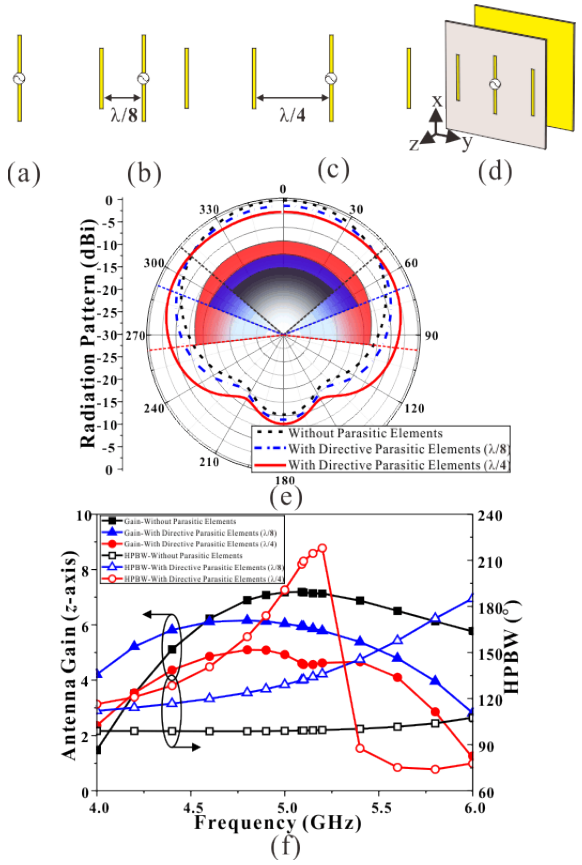


Fig. 2. Performance analysis of the antenna with or without parasitic elements: (a) a dipole antenna without parasitic element, (b) a dipole antenna with two parasitic elements of $\lambda_0/8$, (c) a dipole antenna with two parasitic elements of $\lambda_0/4$, (d) a simulation model for the proposed antenna, (e) radiation patterns and HPBWs with regard to the antenna structure, and (f) antenna gain at the main-beam direction and HPBW with regard to frequency variations.

II. PROPOSED ANTENNA ARRAY WITH A FEEDING CIRCUIT

Based on the linear array antenna theory, the proposed antenna array is based on a two-element dipole antenna array which has broadside beam [8]. Each unit cell antenna operates with the same power and a zero-phase difference. Figure 1 indicates the proposed antenna configuration and its fabricated array prototype. The proposed array consists of an upper substrate, a lower substrate, and two supports in Fig. 1 (a). The upper substrate includes a two-element dipole antenna array and five-parasitic elements with directive and reflective parasitic elements. Parasitic elements are added to widen an elevation beamwidth. To increase the directivity, a ground plane at the top layer on the lower substrate operates as a metallic reflector, and the bottom layer has a feeding network. It consists of a power divider and microstrip lines for feeding the two-element dipole array. To obtain the maximum antenna performance, the separated distance between upper and lower substrates is the height of the supports, and printed microstrip lines and tapered baluns are connected to the radiators. Figures 1 (b) and 1 (c) depict the top view of the upper substrate and antenna supports with a microstrip feeding, respectively. Figures 1 (d) and 1 (e) show the fabricated antenna prototype and the bottom view of the lower substrate, respectively. Table 1 describes the design parameters of the proposed array.

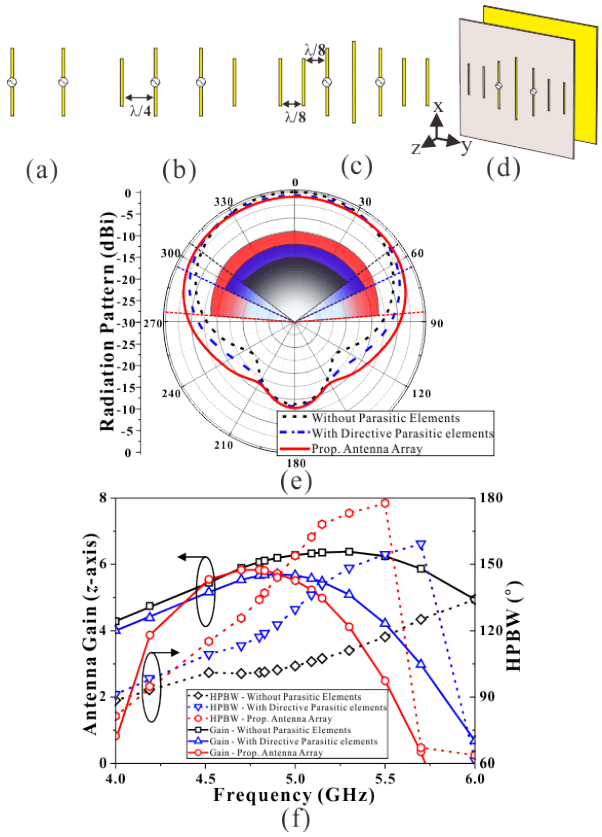


Fig. 3. Comparison between the dipole array with and without parasitic elements: (a) a two-element dipole array without parasitic elements, (b) a two-element dipole array with two parasitic elements, (c) a two-element dipole array with five parasitic elements, (d) a simulation model for the proposed antenna, (e) simulated radiation patterns with regard to the parasitic elements, and (f) simulated antenna gain and HPBW at *z*-axis with regard to the parasitic elements.

Table 1: Design parameters of the proposed array (unit: *mm*)

W_s	W_l	H	D_{p1}	D_r
56	53	20	7.5	17
L_1	L_2	L_3	W_r	W_f
21.4	28	14	0.7	1.78

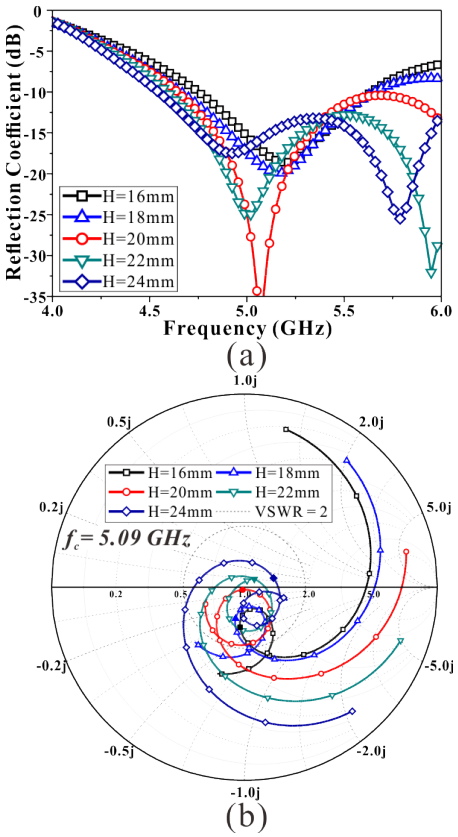


Fig. 4. Simulated impedance variation at operating frequency with regard to H: (a) reflection coefficient variation and (b) impedance variation on Smith chart.

Figure 2 depicts the performance analysis of the different antenna structures with regard to the separated distance between radiators and parasitic elements. Figurees 2 (a)-(c) indicate a dipole antenna, dipole antenna with directive parasitic elements at intervals of λ_0 /8, and a dipole antenna with directive parasitic elements at intervals of λ_0/4, respectively. The simulation has been conducted using an antenna deployment in Fig. 2 (d). It consists of the radiating substrate and a metallic reflector. The simulated radiation patterns and half-power beamwidth (HPBW) at 5.09 GHz are in Fig. 2 (e). As the parasitic elements are far from the dipole antenna, HPBW is increased but the gain is reduced at 0° due to the trade-off between the gain and beamwidth. Similarly, Fig. 2 (f) shows the gain and HPBW with regard to the operating frequency. The antenna gain is decreased and HPBW is increased when the operating frequency increases and the parasitic elements are separated. In particular, the radiated beams are splitted at 5.03 GHz of λ_0/4 deployment.

On the other hand, in order to compare HPBW at the elevation plane with regard to multiple parasitic elements, Figs. 3 (a)–(c) describe the different antenna array deployments with or without multiple parasitic elements. Figure 3 (d) indicates the simulation model for the proposed antenna. Figure 3 (e) shows the simulated radiation patterns of the dipole array with or without two parasitic elements and the proposed antenna array, respectively. The multiply parasitic elements increase the HPBW at elevation plane. The proposed antenna array in Fig. 3 (c) has a wide beamwidth of approximately 162° rather than 106° and 136° at 5.09 GHz in Figs. 3 (a) and 3 (b), respectively. Other comparisons are conducted to compare the antenna gain and HPBW at *z*-axis with regard to the parasitic elements in Fig. 3 (f). As frequency increases, the HPBW also increases. However, the gain is rather reduced at higher than 5.7 GHz due to the beam splitting. Therefore, the proposed antenna array has been optimized for wide elevation angle and the antenna gain at UAV frequency band.

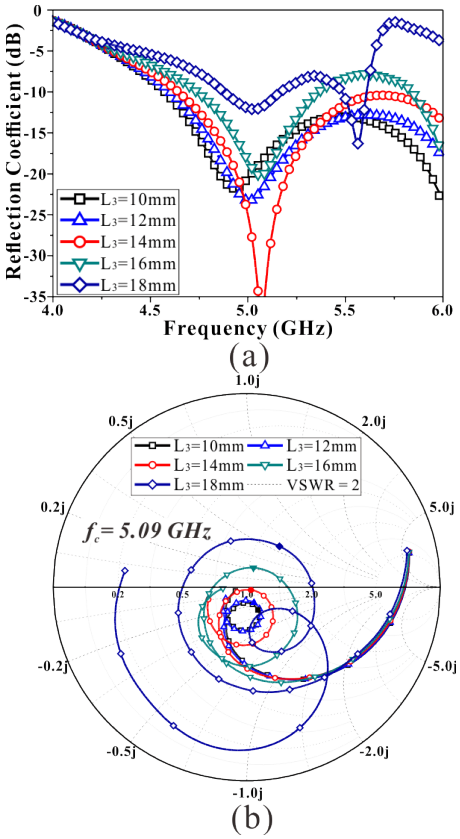


Fig. 5. Simulated impedance variation with regard to L_3: (a) reflection coefficient variation with regard to frequency, and (b) impedance variation on Smith chart.

III. RESULTS AND DISCUSSIONS

The proposed antenna array operates in the UAV communication frequency band from 5.03 GHz to 5.15 GHz, which is assigned to the World Radio communication Conference-12 (WRC) of International Telecommunication Union (ITU). It is simulated and optimized using a commercial full-wave electromagnetic

simulation tool (Microwave Studio 2019 by CST), and designed and fabricated on an RF-35 substrate (ε_r = 3.5, δ = 0.0018) for experimental verifications. To obtain the mutual effect between the upper and lower substrates, the simulated results with regard to the antenna height (H) variations is shown in Fig. 4. The reflection coefficient with regard to the operating frequency is described in Fig. 4 (a). The resonant frequency is increased by reducing *H*. Figure 4 (b) shows the variation of reflection coefficients on Smith chart, and it describes the impedance increases when *H* increases. Similarly, the proposed array is affected by the director length (L_3) in Fig. 5. As the L_3 increases, its resonant frequency and impedance on Smith chart are increased. By considering the antenna gain, radiation patterns with wide beamwidth, and the reflection coefficients, the geometry of the proposed array with multiple parasitic elements is determined on Table 1.

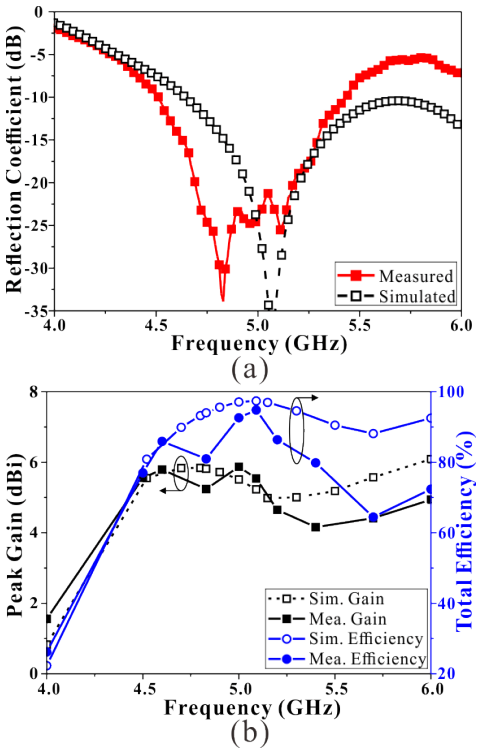


Fig. 6. Simulated and measured antenna performances: (a) reflection coefficient with regard to operating frequency and (b) peak gain and total efficiency with regard to operating frequency.

Figure 6 shows simulated and measured results of reflection coefficient, peak gain, and total efficiency. The measured 10-dB bandwidth is approximately 18% from 4.51 GHz to 5.43 GHz. The difference between simulated and measured results is caused by the implementation error of three-dimensional antenna. The measured peak gain and total efficiencies in Fig. 6 (b) are approximately 5.54 dBi and 95% at 5.09 GHz, respectively. Figure 7 indicates the simulated and measured radiation patterns, which are a good agreement between the simulated and measured results. The HPBW and cross-polarization level of the proposed array are approximately 140° and -23 dB, respectively. Table 2 describe the performance comparison between the previous works and the proposed array. The proposed antenna has relatively wide beam coverage and a high gain.

Table 2: Performance comparison between the previous works and the proposed array

Ref.	f_c (GHz)	Imp. BW (MHz)	Peak Gain (dBi(c))	HPBW (°)	Elect. Size (λ^3)
[1]	1.8	970	8.3	70	0.64×0.64×0.16
[2]	4	600	3.5	221	0.43×0.43×0.24
[3]	1.6	1160	3.4	116	0.35×0.35×0.06
[5]	1.5	1170	4.5	110	0.46×0.46×0.1
[6]	2.55	600	7	120	1.45×1.45×0.12
[13]	2	500	5.5	136	1.58×1.14×0.25
Prop.	**5.09**	**920**	**5.5**	**140**	**0.95×0.9×0.34**

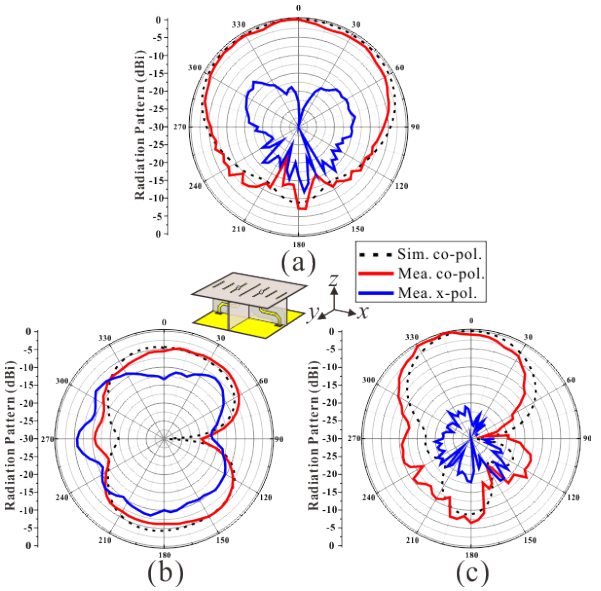


Fig. 7. Simulated and measured radiation patterns with co-polarization and cross-polarization: (a) *zx*-plane, (b) *xy*-plane, and (c) *zy*-plane.

IV. CONCLUSION

A multiply parasitic-coupled, three-dimensional antenna array with wide elevation angle for seamless UAV communication is presented in this paper. Based on the linear array antenna theory, a wide beam-coverage at elevation plane is obtained using a two-element dipole array and parasitic elements. Measured peak gain and

HPBW of the proposed array are approximately 5.5 dBi and 140°, respectively Due to a symmetrical structure, a low-profile configuration, and a wide elevation angle, the proposed antenna array is attractive to various wireless communications systems as well as UAV communications.

ACKNOWLEDGMENT

This work was supported in part by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2017-0-00795, A study on a small antenna system for vehicle supporting wide elevation angle), in part by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIT) (No. 2019R1C1C1008102).

REFERENCES

[1] S. X. Ta and I. Park, "Crossed dipole loaded with magneto-electric dipole for wideband and wide-beam circularly polarized radiation," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 358-361, 2015.

[2] G. Yang, J. Li, S. G. Zhou, and Y. Qi, "A wide-angle E-plane scanning linear array antenna with wide beam elements," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 2923-2926, 2017.

[3] L. Wang, Z. Weng, Y. Jiao, W. Zhang, and C. Zhang, "A low-profile broadband circularly polarized microstrip antenna with wide beamwidth," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, no. 7, pp. 1213-1217, July 2018.

[4] X. Chen, L. Yang, J. Zhao, and G. Fu, "High-efficiency compact circularly polarized microstrip antenna with wide beamwidth for airborne communication," *IEEE Antennas Wireless Propag. Lett.*, vol. 15, pp. 1518-1521, 2016.

[5] W. J. Yang, Y. M. Pan, and S. Y. Zheng, "A low-profile wideband circularly polarized crossed-dipole antenna with wide axial-ratio and gain beamwidths," *IEEE Trans. Antennas Propag*., vol. 66, no. 7, pp. 3346-3353, July 2018.

[6] W. Lin, H. Wong, and R. W. Ziolkowski, "Wideband pattern-reconfigurable antenna with switchable broadside and conical beams," *IEEE Antennas Wireless Propag. Lett*., vol. 16, pp. 2638-2641, 2017.

[7] G. Li and F. Zhang, "A compact broadband and wide beam circularly polarized antenna with shorted vertical plates," *IEEE Access*, vol. 7, pp. 90916-90921, 2019.

[8] C. A. Balanis, *Antenna Theory, Analysis and Design*. New York, Wiley, 1997.

Dong-Geun Seo received the B.S. degree in Electronic Engineering from Gyeongsang National University (GNU), Jinju, South Korea, in 2016, where he is currently pursuing the M.S. degree. His current research interests include near-field wireless power transfer and communication systems, beamforming antenna array and its feeding network for UAV communications, and RFID/IoT sensors.

Wang-Sang Lee received the B.S. degree from Soongsil University, Seoul, South Korea, in 2004, and the M.S. and Ph.D. degrees in Electrical Engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2006 and 2013, respectively.

From 2006 to 2010, he was with the Electromagnetic Compatibility Technology Center, Digital Industry Division, Korea Testing Laboratory (KTL), Ansan-si, South Korea, where he was involved in the international standardization for radio frequency identification (RFID) and photovoltaic systems as well as electromagnetic interference (EMI)/EMC analysis, modeling, and measurements for information technology devices. In 2013, he joined the Korea Railroad Research Institute (KRRI), Uiwang-si, South Korea, as a Senior Researcher, where he was involved in the position detection for high-speed railroad systems and microwave heating for low-vibration rapid tunnel excavation system. Since 2014, he has been an Associate Professor with the Department of Electronic Engineering, Gyeongsang National University (GNU), Jinju, South Korea. From 2018 to 2019, he was a Visiting Scholar with the ATHENA Group, Georgia Institute of Technology, Atlanta, GA, USA. His current research interests include near- and far-field wireless power and data communications systems, RF/microwave antenna, circuit, and system design, RFID/Internet-of-Things (IoT) sensors, and EMI/EMC.

Lee is a member of IEC/ISO JTC1/SC31, KIEES, IEIE, and KSR. He was a recipient of the Best Paper Award at IEEE RFID in 2013, the Kim Choong-Ki Award: EE Top Research Achievement Award by the Department of Electrical Engineering, KAIST, in 2013, the Best Ph.D. Dissertation Award: the Department of Electrical Engineering, KAIST, in 2014, the Young Researcher Award by KIEES in 2017.

Compact Substrate Integrated Waveguide Band-pass Rat-race Couplers Based on Mixed Shape Cavity with Flexible Port Topology

Zhigang Zhang, Yong Fan, and Yonghong Zhang

Fundamental Science on Extreme High Frequency Key Laboratory
University of Electronic Science and Technology of China, Chengdu 611731, China
freemanzzg@163.com, yfan@uestc.edu.cn, zhangyhh@uestc.edu.cn

Abstract ─ A new type of compact band-pass rat-race couplers with flexible port topology based on mixed shape substrate integrated waveguide (SIW) cavity is first proposed in this paper. Moreover, composite right/left-handed (CRLH) SIW unit loaded on the common wall between cavities is used to achieve negative coupling structure. The diameter of the vias on the common wall related CRLH-SIW unit is also used to adjust the electric coupling strength and increase the coupling bandwidth. The detailed analysis and the design method have been introduced to realize a filtering rat-race coupler based on substrate integrated fan-shaped cavity (SIFC) and rectangular cavity (SIRC). In particular, the combination of mixed shape resonators and port location in rectangular resonators can be selected according to the requirement of port angle interval. Compared with other filtering couplers, the proposed designs exhibit good filtering responses, high *Q* factor, better isolation, multiple port angle intervals, amplitude balance, as well as 0° and 180° phase differences.

Index Terms ─ Band-pass coupler, flexible port topology, mixed shape cavity, negative coupling structure, substrate integrated fan-shaped cavity (SIFC).

I. INTRODUCTION

Rat-race couplers are essential components in transceivers for microwave communication systems. Miniaturized couplers with high performance are usually required in many applications. As a promising candidate for modern wireless transceiver systems, SIW technology [1-8] is beneficial to the realization of miniaturized couplers and filters due to the fact that it provides an excellent trade-off between waveguide and planar technologies. In [2], synthesis and design techniques of dual-band filters are proposed. A compact multilayer dual-mode filter based on the substrate integrated circular cavity (SICC) is developed in [6].

To further reduce size, a single device integrated with different functionalities, such as filtering coupler [9-21], has been attracting increasing attention. It also recommends an effective way to avoid the performance degradation due to a cascade connection of two individual components. Bandpass 90° and 180° directional couplers with coupled resonators have been first proposed in [9]. A compact filtering 180° hybrid is presented in [12]. Moreover, it is noteworthy that the coupler topologies also have an important impact on the miniaturization of the microwave system. Therefore, the port of the components should be placed at proper positions to meet the requirement of system topology and interconnect.

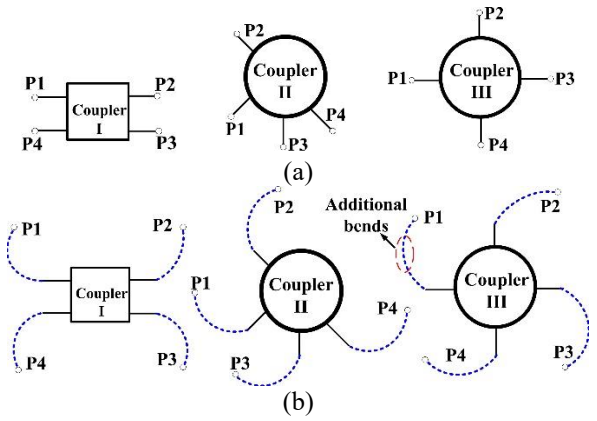


Fig. 1. (a) Different types of topologies of existed couplers, and (b) additional bends are employed to match with the desired transmission direction.

Recently, as shown in Fig. 1 (a), the existed couplers have several different types of topologies. 1) Two ports are placed at one side, and others are placed at the opposite side [9-10],[17],[20]; 2) the distance between each two ports is determined by the required electric length [12-15]; 3) ports are mutually orthogonal [18-19]. By using degenerate orthogonal modes of square SIW cavity, a crossover with bandpass responses is designed in [18]. Four TE_{101}-mode square SIW cavities based on multilayer coupling structures are utilized to construct the filtering rat-race coupler [20]. A SIW filtering rat-race coupler based on TE_{102} and TE_{201} orthogonal degenerate modes is proposed in [21].

However, these types of topologies are not enough

Submitted On: December 23, 2019
Accepted On: February 28, 2020

in the high-density circuit design. As such, additional bends (Fig. 1 (b)) will be employed to modify their line direction to meet the system requirement, which leads to a large circuit area, a complex configuration and unwanted loss, electromagnetic compatibility (EMC) problem, especially leakage from the discontinuity. Therefore, it is expected that the port angle interval can be selected flexibly according to the topological requirements without introducing additional L-bends.

In this paper, the mixed shape SIW cavities are first proposed to design band-pass rat-race couplers with flexible port topology. Specifically, by using mixed shape cavities, the structure of filtering coupler can be reasonably constructed according to the topology requirements while keeping good performance. With flexible ports arrangement and resonators combination, the coupler's topology is easy to match with the desired transmission direction, which means that there is no need to add additional L-bend and occupy additional area. What's unique about the analysis process of coupler is that the band-pass coupler is divided into two filtering power dividers, then each power divider is divided into two filters respectively, and the influence of the angle intervals between ports on *Q*e is considered in the design of coupler. Afterwards, the equivalent circuit and coupling matrix method are used to evaluate the initial value of design parameters accurately according to the specifications, which is beneficial to accelerate the later optimization design process. Furthermore, the demand of circuit topology has been integrated into design considerations for the first time, which is helpful to meet the requirement of system interconnect and the design of miniaturized structures.

II. ANALYSIS AND DESIGN

A. Configuration

Several four-element coupler topologies are depicted in Figs. 2 and 3 to demonstrate different ways to achieve different port angular intervals.

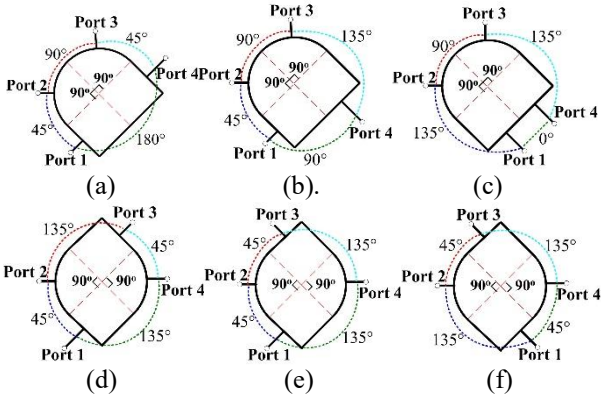


Fig. 2. (a)~(f) Examples of available angular intervals for Type A filtering coupler.

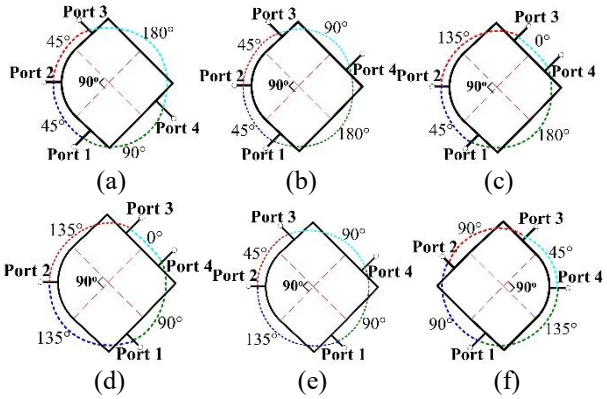


Fig. 3. Examples of available angular intervals for Type B filtering coupler.

The shape of the resonator is very important to obtain the desired port direction while keeping good performance. Rectangular resonators [22-29] are easy to realize 0°, 90° and 180° interval between adjacent ports. Fan-shaped resonators may be a universal configuration. Furthermore, these different shape resonators can be mixed to achieve type A and B couplers with flexible port intervals. The combination of resonators with different shapes and the choice of port location in rectangular resonators are the key factors to construct various angular intervals for input/output ports. Substrate integrated fan-shaped cavities (SIFC) and rectangular cavities (SIRC) are utilized as basic resonance cells of the proposed filtering couplers.

B. Analysis of coupler

The topology of the filtering coupler is shown in Fig. 4 (a). As seen, the rat-race coupler has following two working states. When signals are injected from port 1, the output signals from ports 2 and 4 are in-phase, equal power allocation (M_{21}=M_{41}), just as an in-phase power divider should do. If port 3 is excited, the out-of-phase responses are obtained in ports 2 and 4 (-M_{23}=M_{43}). In this case, the coupler can be seen as an out-of-phase power divider.

As depicted in Figs. 4 (a) and (b), in these two working states, the bandpass coupler is equivalent to an in-phase filtering power divider and an out-of-phase filtering power divider, respectively. Moreover, divider I and II shown in Figs. 4 (b) and (c) are both designed with 3-dB power split ability and have same passband characteristics but quite different phase characteristic. Therefore, the coupling coefficients between resonators of dividers have the following relationship:

$$M_{21}=M_{41}=-M_{23}=M_{43}. \quad (1)$$

In addition, each power divider can be divided into two second-order bandpass filters which have the same operating frequency and passband characteristics, as shown in Fig. 4 (d) ($M_{21,\,BF}$=$M_{41,\,BF}$, $M_{2L,\,BF}$=$M_{4L',\,BF}$).

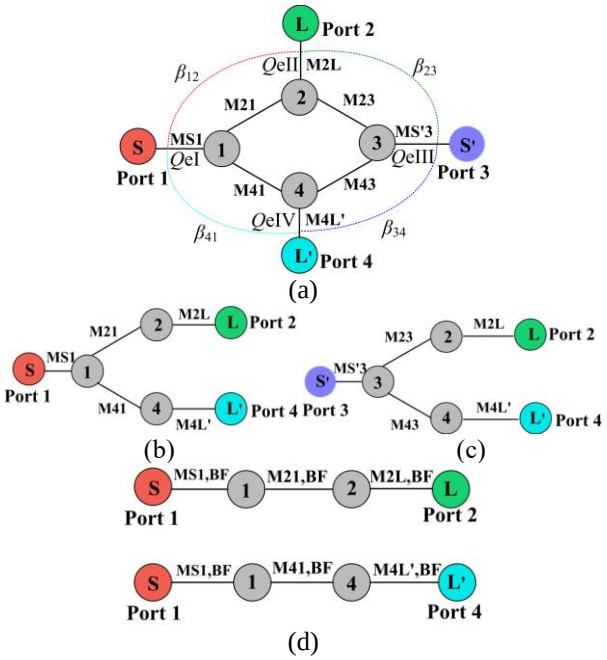


Fig. 4. (a) Topology of the filtering coupler, (b) topology of the filtering in-phase power divider, (c) topology of the filtering out-of-phase power divider, and (d) topology of two second-order filters. Where, β_{12}, β_{23}, β_{34}, and β_{41} are angular intervals for each port, respectively.

Thus, the filtering rat-race coupler can be divided into four bandpss filters. Our strategy is to separately design these four components, and then combine them into one coupler. The coupling matrix of this second-order coupled-resonator band-pass filter is expressed as:

$$m_{N+2}=\begin{array}{c} \\ S \\ 1 \\ 2 \\ L \end{array}\begin{array}{c} \begin{array}{cccc} S & 1 & 2 & L \end{array} \\ \begin{bmatrix} 0 & m_{S1,BF} & 0 & 0 \\ m_{1S,BF} & 0 & m_{12,BF} & 0 \\ 0 & m_{21,BF} & 0 & m_{2L,BF} \\ 0 & 0 & m_{L2,BF} & 0 \end{bmatrix} \end{array}. \quad (2)$$

And the normalized input impedance of the filtering power divider in Fig. 4 (b) is required to be the same as matrix (2). Thus, the coupling coefficients for the filtering power divider topology in Fig. 4 (b) are determined as:

$$M_{S1,BF}=M_{S1}, \quad (3a)$$

$$M_{2L}=M_{4L}=M_{2L,BF}=M_{4L',BF}. \quad (3b)$$

Figure 5 is the equivalent circuit of the filtering coupler. As seen, four parallel *LC* resonators are used to model these four SIW resonators. The in-phase couplings of "resonator 1-to-2" and "resonator 1-to-4" are modeled by two interconnected -90° *J*-inverters (admittance inverters). The 180° out-of-phase couplings of "resonator 3-to-2" and "resonator 3-to-4" are modeled by one +90° and one -90° *J*-inverters, respectively. On the other hand, when port 1 is excited, there are two paths for a signal to reach port 3, one is resonator 1-to-2-to-3 and the other is resonator 1-to-4-to-3. On the first path, there are one "*J*-90°" and one "*J*+90°" inverters so the signal accumulates the phase of 0°. On the second path, there are two "*J*-90°"inverters which lead to 180° accumulated phase. These two signals cancel out each other and result in zero voltage on port3. Thus, isolation between ports 1 and 3 is achieved.

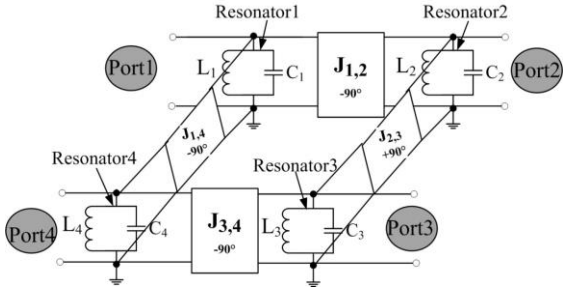


Fig. 5. Equivalent lumped circuit model of the filtering rat-race coupler.

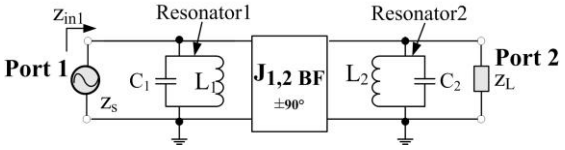


Fig. 6. Equivalent circuit of the coupled-resonator bandpass filter.

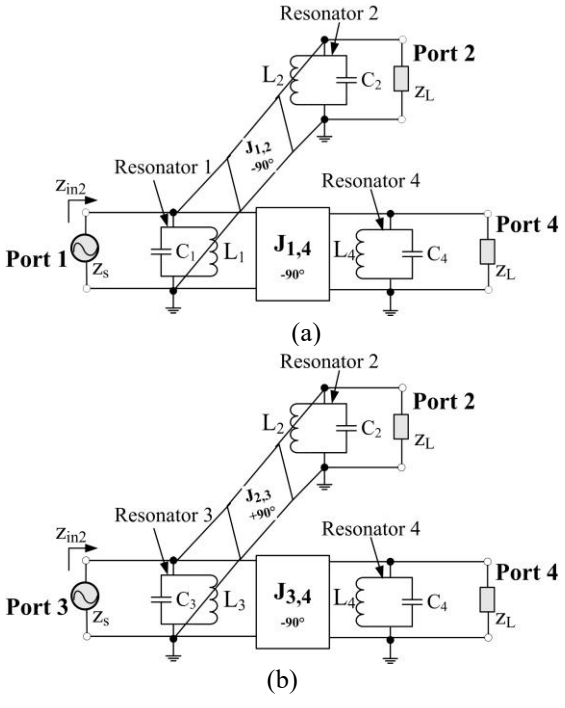


Fig. 7. Equivalent circuit of the filtering power divider: (a) in phase power divider, and (b) out-of-phase power divider.

The cross couplings between "resonator 1-to-3" and "resonator 2-to-4" are neglected in this model because the distances between them are relatively longer. Since the proposed circuit is reciprocal, it can also be excited from port 2 or port 4.Based on the analysis of the above-mentioned topology, the filtering coupler can be seen as consisting of four coupled-resonator filters. The equivalent circuit of the two-order coupled-resonator filter is shown in Fig. 6. The input impedance of this filter is calculated as:

$$z_{in1} = z_s + \frac{j\omega L_1}{1-\omega^2 L_1 C_1} + \frac{\left(j\omega J_{1,2BF}\right)^2}{\frac{j\omega L_2}{1-\omega^2 L_2 C_2} + z_L}. \quad (4)$$

According to the topology of the filtering coupler, the Equivalent circuit of the filtering divider is depicted in Fig. 7. The equivalent circuit of the filtering power divider in Fig. 7 can be realized by adding the nodes "*J*-90°"inverters and resonator 4 to the circuit in Fig. 6. The input impedance of equivalent circuit in Fig. 7 is calculated as:

$$z_{in2} = z_s + \frac{j\omega L_1}{1-\omega^2 L_1 C_1} + \frac{\left(j\omega J_{1,2}\right)^2}{\frac{j\omega L_2}{1-\omega^2 L_2 C_2} + z_L} + \frac{\left(j\omega J_{1,4}\right)^2}{\frac{j\omega L_4}{1-\omega^2 L_4 C_4} + z_L}. \quad (5)$$

In Fig. 5, four resonators are to be designed with the same external quality factor and resonant frequency (L_1= L_2= L_3= L_4=L_0, C_1= C_2= C_3= C_4=C_0). For a 3-dB divider (J_{12}= J_{14}), the input impedance is also given by:

$$z_{in2} = z_s + \frac{j\omega L_1}{1-\omega^2 L_1 C_1} + \frac{\left(j\omega\sqrt{2} J_{1,2}\right)^2}{\frac{j\omega L_2}{1-\omega^2 L_2 C_2} + z_L}. \quad (6)$$

In order to have the same return-loss performance for these two circuits, z_{in1} should be equal to z_{in2}. Based on Eqs. (4) and (6), we have:

$$J_{1,2} = J_{1,4} = \frac{J_{1,2BF}}{\sqrt{2}}. \quad (7)$$

Moreover, the coupling coefficients between resonators [31] are:

$$M_{12} = M_{14} = \frac{J_{1,2}}{\sqrt{\omega_1 C_1 \omega_2 C_2}} = \frac{J_{1,4}}{\sqrt{\omega_1 C_1 \omega_4 C_4}}, \quad (8a)$$

$$M_{12,BF} = M_{21,BF} = \frac{J_{1,2BF}}{\sqrt{\omega_1 C_1 \omega_2 C_2}}, \quad (8b)$$

From Eqs. (7) and (8), the relationship between the coupling coefficients of the filter and the divider is as follows:

$$M_{12} = M_{14} = \frac{M_{12,BF}}{\sqrt{2}}. \quad (9)$$

Based on (1) ~ (3) and (9), the coupling matrix for the filtering power divider topology in Fig. 4 (b) is determined as:

$$M_{S1,BF} = M_{S1}, \quad (10a)$$

$$M_{2L} = M_{4L'} = M_{2L,BF} = M_{4L',BF}, \quad (10b)$$

$$M_{21} = M_{41} = -M_{23} = M_{43} = \frac{M_{21,BF}}{\sqrt{2}} = \frac{M_{41,BF}}{\sqrt{2}}. \quad (10c)$$

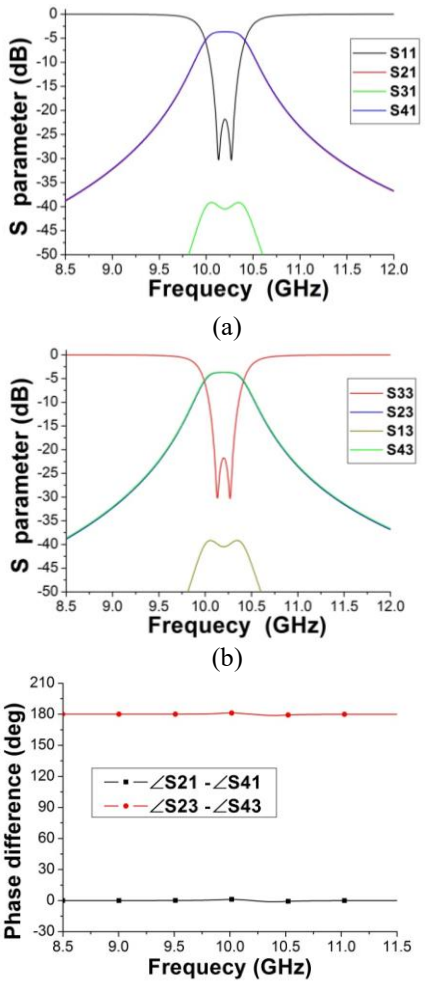


Fig. 8. Circuit simulation of the proposed filtering rat-race coupler: (a) Port1 excited, (b) Port3 excited, and (c) Phase differences.

The required normalized coupling coefficient (m) and external quality factors (Q_e) for the filtering power divider can be calculated by:

$$m_{12} = \frac{M_{21}}{FBW} = \frac{M_{41}}{FBW}, m_{12,BF} = \frac{M_{21,BF}}{FBW}. \quad (11)$$

$$Q_e = \frac{FBW}{M_{S1}^2} = \frac{FBW}{M_{2L}^2} = \frac{1}{FBW \times m_{S1}^2}. \quad (12)$$

Generally, external quality factor (Qe) is related to the following three parameters: the length of feeding slot (Lgap), the width of external coupling aperture (Wc), port angular interval (β). So, the Qe corresponding to

each port of coupler is denoted as: *Q*eI (Wc1, Lgap, β_{12}), *Q*eII (Wc2, Lgap, β_{23}), *Q*eIII (Wc3, Lgap, β_{34}), *Q*eIV (Wc4, Lgap, β_{41}), respectively. In order to obtain the same return loss and passband characteristics on each port, we have:

$$Q\text{eI}= Q\text{eII}= Q\text{eIII}= Q\text{eIV}=Q\text{e}. \quad (13)$$

Since the equivalent circuit in Fig. 6 is a second-order coupled-resonator bandpass filter, it can be design by the conventional synthesis technology [31]. Take the second-order Chebyshev equal-ripple response for example, with the center frequency f_0, ripple level (L_a) (dB), and fractional bandwidth (FBW), and the port admittance Y_0, the element values can be calculated:

$$L=\frac{\sinh(u/4)}{\sqrt{2}}\frac{FBW}{2\pi f_0 Y_0}, \quad (14)$$

$$C=\frac{\sqrt{2}}{\sinh(u/4)}\frac{Y_0}{2\pi f_0\times FBW}, \quad (15)$$

$$J_{1,2BF}=\frac{1}{\sqrt{2}}Y_0\coth(u/4), \quad (16)$$

$$u=\ln[\coth(\frac{\ln 10}{40}\times L_a(dB))]. \quad (17)$$

In this design, a 50-Ω (Y_0=0.02S) filtering rat-race coupler is centered at 10.2 GHz with second-order Chebyshev 0.05-dB equal-ripple response [L_a=0.05(dB)] and 2.6% fractional bandwidth (FBW). Based on above analysis, the filtering coupler can be seen as consisting of four coupled-resonator filters. The parameter values of the equivalent circuits in Fig. 6 and Fig. 7 are then evaluated using (7)–(9),(14)–(17): u=5.85058, L_1= L_2= L_3= L_4=L_0=0.0293nH, C_1= C_2= C_3= C_4=C_0=8.309pF, $J_{1,2\,BF}$=0.01575S, Y_0=0.02S. $J_{1,2}$=$J_{1,4}$=-$J_{2,3}$=$J_{3,4}$=0.0111369S. After determining these parameters, the circuit structure can be constructed to obtain the desired responses.

Figures 8 (a) and (b) show the circuit simulation of the proposed filtering rat-race coupler under the in-phase and out-of-phase operation, respectively. As seen, good filtering responses and amplitude balance can be obtained. Since the whole structure is symmetric, good phase imbalance of less than 1° between ports 2 and 4 is achieved, as shown in Fig. 8 (c).

C. Design consideration

The resonant frequency of mode for circular cavity with solid wall can be calculated by [24],[33]:

$$f_{mpn}=\begin{cases}\dfrac{c}{2\pi\sqrt{\mu_r\varepsilon_r}}\sqrt{\left(\dfrac{\mu'_{mn}}{R}\right)^2+\left(\dfrac{p\pi}{\Delta h}\right)^2} & TE_{mpn}\\ \dfrac{c}{2\pi\sqrt{\mu_r\varepsilon_r}}\sqrt{\left(\dfrac{\mu_{mn}}{R}\right)^2+\left(\dfrac{p\pi}{\Delta h}\right)^2} & TM_{mpn}\end{cases}, \quad (18)$$

where μ_r and ε_r are relative permeability and permittivity of the filling material, μ_{mn} and μ'_{mn} are the nth roots of mth Bessel function of the first kind and its derivative, R is the radius of circular cavity, Δh is the height of the of circular cavity ,and c is the speed of light in free space. According to Eq. (18) and by means of the least square method, the resonant frequency of the TM_{101} mode for SIFC can be calculated by the following formula:

$$f_{101}=\frac{0.383c}{\dfrac{1}{b_\theta}R_{eff}\sqrt{\mu_r\varepsilon_r}},\ R_{eff}=R-\frac{D^2}{0.95p}. \quad (19)$$

Where, θ is the central angle of SIFC, R_{eff} is the equivalent radius of the fan-shaped cavity. D and p are the diameter of metallized via-holes and center-to-center pitch between two adjacent via-holes. b_θ is related to the central angle of a fan-shaped resonator. When θ=45°, 60°, 90°, 120°, b_θ is approximately equal to 3.2, 2.7, 2.1, 1.8, respectively.

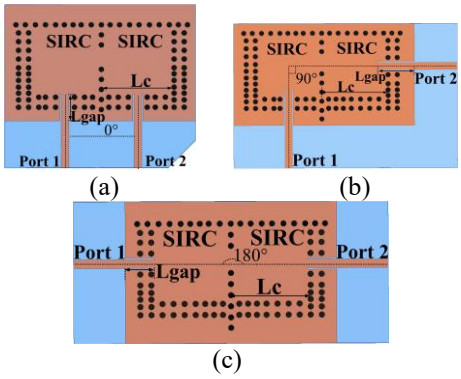


Fig. 9. Several port angular intervals of two SIRCs: (a) 0°, (b) 90°, (c) 180°. Where, *L*c is the side length of the SIRC, *L*gap is the length of feeding slot.

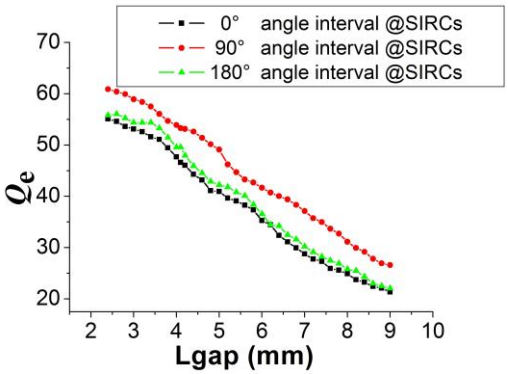


Fig. 10. Extracted *Q*e curves corresponding to different port angle intervals of two SIRCs.

Figure 9 shows the combination of resonators facilitates the acquisition of different port angle intervals and the construction of different coupler topologies. At the same time, however, external quality factor (*Q*e) is also affected by the direction of ports. This means that besides the influence of external coupling parameters on *Q*e, port angle intervals also play an important role.

Therefore, numerical analysis is carried out on SIRCs with microstip lines as its input/output to determine the external quality factor using Ansoft's HFSS. The external coupling is controlled by changing the length of feeding slot (*L*gap) with a fixed feeding slot width.

Figure 10 depicts extracted *Q*e curves versus the length of feeding slot (*L*gap). As we can see, there is little difference between *Q*e variations at 0°and 180° port angular intervals in two SIRCs. When port angular interval is 90°, the value of corresponding *Q*e is larger than that of the previous two cases.

III. FILTERING COUPLER TYPE A

A. Coupler structure

As shown in Fig. 11 (a), the filtering Type A coupler consists of two fan-shaped SIW cavities and two square cavities. The coupling between resonators I and II, I and IV, III and IV is magnetic coupling, which is realized by the iris between the common narrow walls. However, electrical coupling occurs between resonators II and III due to composite right/left-handed (CRLH) transmission line loaded on the common walls.

As shown in Fig. 11 (b), CRLH-SIW unit is achieved by introducing an interdigital slot etched on the metal surface of common post wall of SIW cavities. It should be noted that CRLH-SIW unit is used to achieve negative coupling structure, which decreases the resonant frequency of the corresponding SIW cavities.

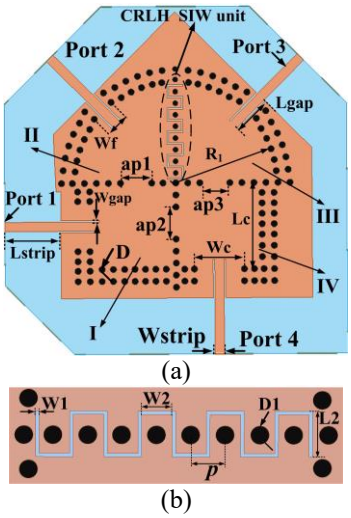


Fig. 11. The proposed Type A filtering coupler: (a) top view and (b) CRLH-SIW unit.

Figure 12 shows the CRLH is loaded on the common walls oure the SIFCs, SIRCs and mixed SIW cavity respectively, resulting in decrease of the central frequency. The main reason is that the negative order resonance is excited in corresponding SIW cavities, which leads to decrease of the cut-off frequency of the waveguide [25]. In addition, the central frequency offset of the SIFCs is larger than that of other cases due to the loading of CRLH, as can be seen from Fig. 12.

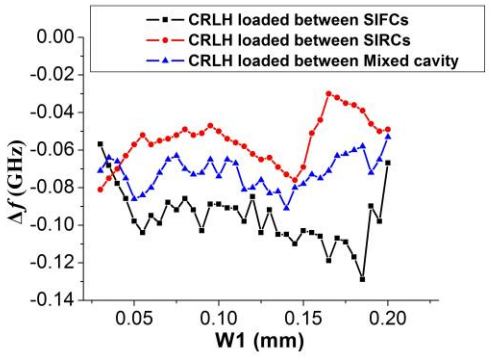


Fig. 12. Variation of the central frequency of SIW cavities with CRLH loaded. Where, $\Delta f=f_1-f_0$, f_1 and f_0 are the central frequency of cavities with and without CRLH loaded, respectively.

B. Design

In our design, the desired passband is centered at 10.2 GHz with the 2.6% fractional bandwidth (FBW) of 19-dB equal-ripple return loss. Based on the advanced coupling matrix synthesis method in [32], the initial normalized coupling matrix of corresponding bandpss filter (BPF) can be synthesized as:

$$m_{N+2}=\begin{array}{c} \\ S \\ 1 \\ 2 \\ L \end{array}\begin{array}{c} \begin{array}{cccc} S & 1 & 2 & L \end{array} \\ \begin{bmatrix} 0 & 1.18591 & 0 & 0 \\ 1.18591 & 0 & 1.57423 & 0 \\ 0 & 1.57423 & 0 & 1.18591 \\ 0 & 0 & 1.18591 & 0 \end{bmatrix} \end{array}. \quad (20)$$

From (7)–(12), (20) the desired parameters of the filtering power dividers can be calculated as follows: $M_{21}= M_{41}= -M_{23}= M_{43}=0.02894$, $Q_{e,S1}=27.3463$, $Q_{e,2L}= Q_{e,4L}=27.3463$. The external quality factor *Q*e is calculated by [31]

$$Q_e=\frac{2f_0}{\Delta f_{3dB}}, \quad (21)$$

where f_0 is the frequency at which S21 reaches its maximum value and Δf_{3dB} is the 3-dB bandwidth for which S21 is reduced by 3 dB from its maximum value.

Figure 13 shows the external quality factor versus the width of the external coupling aperture (*W*c) and the length of feeding slot (*L*gap). It can be seen that the larger the feeding slots length and the external coupling aperture width, the smaller the external quality factor. Based on the previous analysis, the filtering rat-race coupler can be divided into four bandpss filters. The adjustment of external quality factor (*Q*eI, *Q*eII) is realized by changing the value of Lgap, as depicted in

Fig. 13 (a). Then, the desired Qe (QeII = QeIII = Qe) can be achieved and the initial value of Lgap can also be determined. Due to different port angle intervals, the values of QeI and QeIV are greater than that of QeII and QeIII, as can be clearly seen from Fig. 13 (a). Therefore, another external coupling parameter (Wc) needs to be adjusted to achieve the same Qe as Port2 and 3. Figure 13 (b) gives the extracted curve of QeI versus the coupling aperture width Wc, which is determined to meet the requirements with QeI=QeIV=Qe. Thus, the initial value of Wc can be evaluated. Figure 14 illustrates the extracted curves of M_{21} and M_{41}, which versus the coupling iris ap1 and ap2. As can be seen, when the width of the coupling iris increases, coupling coefficients increases accordingly. The coupling coefficients versus the interdigital slot widths (W1) and diameter of the vias (D1) are plotted in Fig. 14. The diameter of the vias on the common wall related CRLH-SIW unit is used to adjust the electric coupling strength and increase the coupling bandwidth. Obviously, smaller W1 and D1 correspond to the wider bandwidth. In summary, the design procedure of the proposed filtering coupler is listed as follows.

Firstly, according to the requirement of the topology structure, the combination of resonators with different shapes can be selected in Fig. 2. Secondly, the resonant frequency of the SIFC is calculated by formula (19), to meet the required center frequency f_0. Thirdly, a coupling matrix of a second-order BPF is synthesized according to the desired center frequency f_0 and the fractional bandwidth (FBW). Fourthly, according to formula (7) ~ (12), the coupling matrix and Qe of the corresponding filtering power divider are obtained. Moreover, internal coupling parameters (ap1, ap2, ap3) and external coupling parameters (Wc, Lgap) are tuned to meet desired values of coupling coefficients and external quality factor, respectively. Finally, fine tuning of the whole structure is performed to realize good filtering coupler performance.

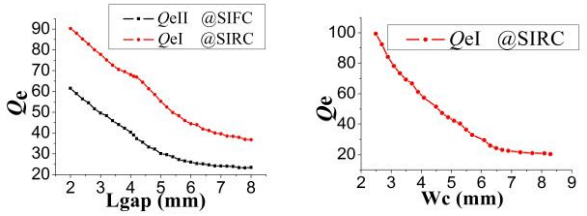


Fig. 13. (a) Simulated external quality factor Qe change with Lgap, Wc=4mm. (b) Simulated QeI change with Wc, Lgap=5mm. Where, QeI=QeIV, QeII=QeIII correspond to port 1, port4 and port2, port3, respectively.

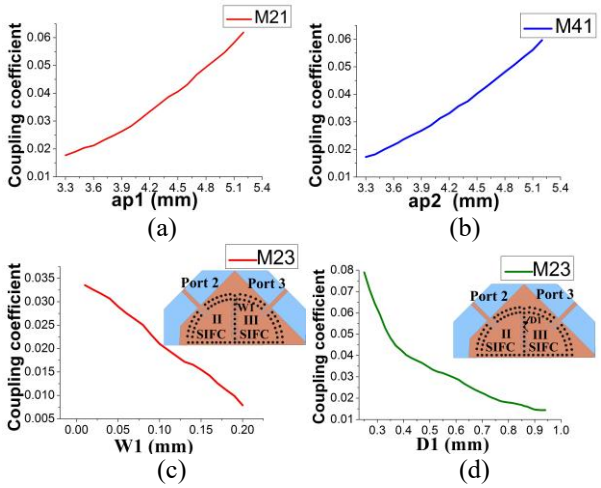


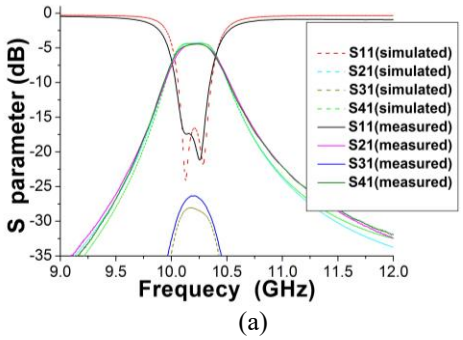
Fig. 14. Coupling coefficients versus the coupling iris: (a) M21 versus ap1, (b) M41 versus ap2, (c) M23 versus W1, and (d) M23 versus D1.

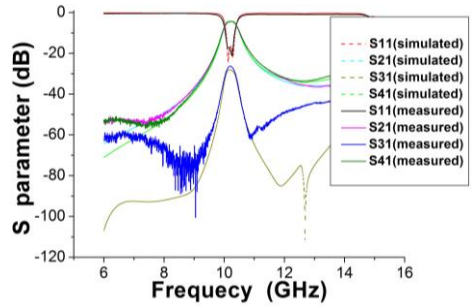
C. Experiment

After optimization implemented by HFSS, the geometry parameters of the proposed coupler are chosen as follows (all in mm): D=0.8, p=1.5, ε_r=3.5, L2=2, W1=0.12, W2=1.35, Wgap=0.2, Lgap=5.1, Wc=6.25, Wstrip =1.15, Lstrip=7, ap1=3.9, ap2=3.95, ap3=3.75, D1=0.59, R_1= 12.63, Lc=10.75, Wf=3.86.

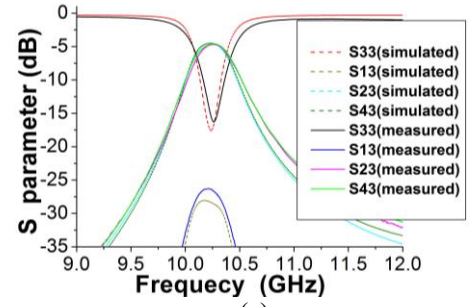
To verify the above method ,the proposed coupler was designed and fabricated on a substrate with thickness of 0.508 mm, relative dielectric constant of 3.5 and dielectric loss tangent 0.0018 (at 10 GHz). The measurement is accomplished by using the Agilent N5244A network analyzer.

Figures 15 (a), (b) shows the simulated and measured S-parameters under the in-phase operation. The measured passband is centered at 10.23 GHz with the 1-dB FBW of 3.5% .The in-band return loss is better than 17.35 dB. The minimum insertion losses including the 3-dB equal power division loss are (3+1.3) and (3+1.35) dB, with the amplitude imbalance of 0.1 dB.

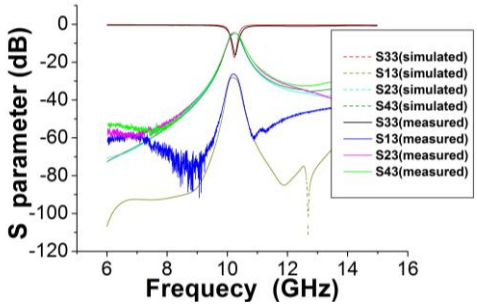




(b)



(c)



(d)

Fig. 15. Simulated and measured results of the fabricated type A coupler with two SIRCs and two SIFCs (a) and (b) Port 1 is excited, (c) and (d) Port 3 is excited.

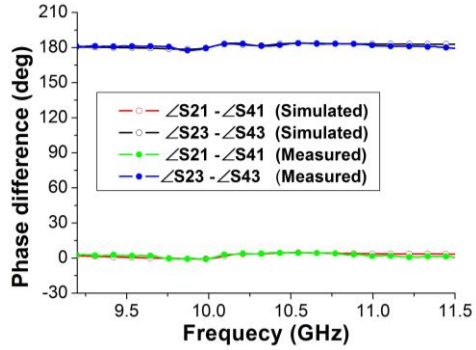


Fig. 16. Phase differences of the fabricated type A coupler.

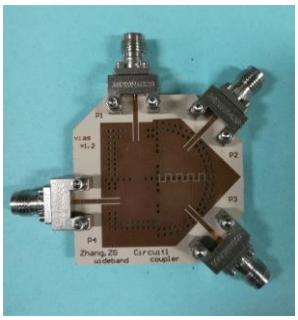


Fig. 17. Photograph of the fabricated filtering type A coupler with two SIRCs and two SIFCs.

The simulated and measured S-parameters under the out-of-phase operation are shown in Figs. 15 (c), (d). The 1-dB FBW of 2.75% is a bit lower than that of the in-phase operation of 3.5%, which is due to the fabrication errors. The insertion losses at the center frequency are (3+1.4) dB and (3+1.45) dB, with the difference of less than 0.1 dB. The measured return loss of port 3 is better than 16.3 dB. In Fig. 15, the isolation between ports 1 and 3 within the passband is better than 26.36 dB, showing high isolation.

When signals are injected from port 1 and port 3, the measured in-band phase differences between two output ports are nearly 0° and 180°, respectively, with the variation of less than 3°, as depicted in Fig. 16. In general, the measured results show good performance of bandpass filtering responses, equal power. Figure 17 is the photograph of the fabricated filtering type A coupler based on two SIRCs and two SIFCs.

IV. FILTERING COUPLER TYPE B

A. Coupler structure

As shown in Fig. 18, the filtering Type B coupler is composed of three square cavities (resonators I~III) and one fan-shaped cavities (resonators IV). In these figures, R_1 is the radius of the SIFC, ap1, ap2 and ap3 are the widths of the coupling iris between the common narrow walls of SIW cavities, *W*c1 and *W*c2 are the width of the external coupling aperture. It should be noted that CRLH-SIW unit is used to achieve negative coupling structure, which decreases the resonant frequency of the corresponding SIW cavities. Therefore, the size of resonators II and III will be smaller than the resonator I (*L*c1<*L*c2).

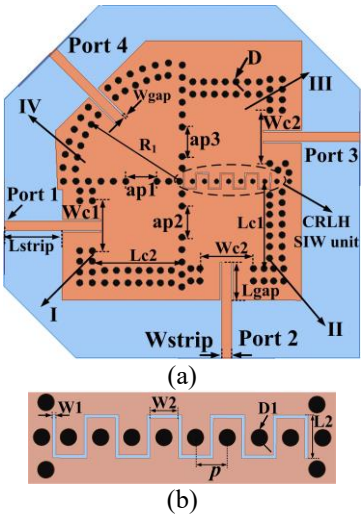
B. Design

In our design, the desired passband is centered at 10.28 GHz with the 1.5% fractional bandwidth (FBW)

of 20-dB equal-ripple return loss. The initial normalized coupling matrix of corresponding BPF can be synthesized as:

$$m_{N+2} = \begin{array}{c} \\ S \\ 1 \\ 2 \\ L \end{array} \begin{array}{c} \begin{array}{cccc} S & 1 & 2 & L \end{array} \\ \begin{bmatrix} 0 & 1.22474 & 0 & 0 \\ 1.22474 & 0 & 1.65831 & 0 \\ 0 & 1.65831 & 0 & 1.22474 \\ 0 & 0 & 1.22474 & 0 \end{bmatrix} \end{array} \cdot (22)$$

From (7)–(12), (22) the desired parameters of the filtering power dividers can be calculated as follows: $M_{21}= M_{41}=-M_{23}=M_{43}=0.017589$, $Qe,_{S1}=44.4453$, $Qe,_{2L}= Qe,_{4L}=44.4453$. In general, the design procedure of the proposed filtering coupler Type B is listed as follows. First, the combination of resonators with different shapes can be selected in Fig. 3. Second, the resonant frequency of the SIFC is calculated by formula (19). Third, a coupling matrix of a second-order BPF is synthesized according to the desired center frequency f_0 and the FBW. Fourth, according to formula (10) ~ (12), the coupling matrix and Qe of the corresponding filtering power divider are obtained. Moreover, internal and external coupling parameters are tuned to meet desired values of coupling coefficients and external quality factor, respectively. Finally, the optimization is implemented by HFSS and good filtering rat-race coupler performance can be realized.



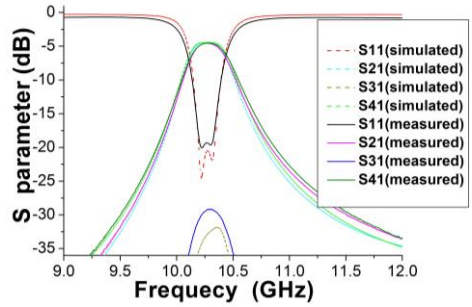
(a)

(b)

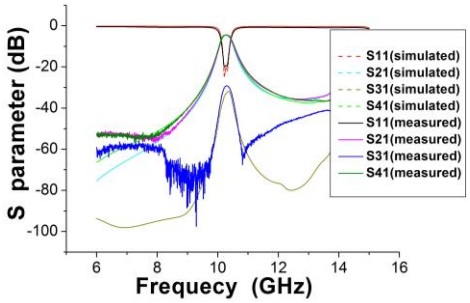
Fig. 18. The proposed filtering Type B coupler with three SIRCs and a SIFC: (a) Top view and (b) CRLH-SIW unit.

C. Experiment

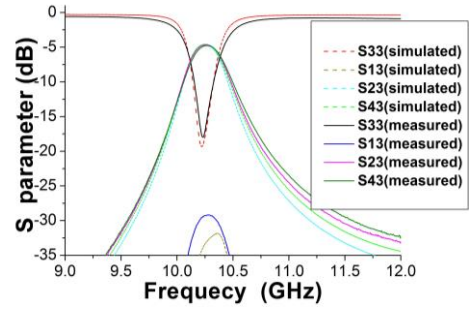
To further verify the proposed method, the filtering rat-race type B coupler based on three SIRCs and a SIFC is implemented. The dimensions of the coupler are chosen as follows (all in mm): $L2$=2, $W1$=0.13, $W2$=1.3, Wgap=0.2, Lgap=4.5, Wc1=5.93, Wc2 =6.33, Wstrip =1.15, Lstrip=7, ap1=3.65, ap2=3.65, ap3=3.55, $D1$=0.58, $R1$= 12.9, Lc1=10.53, Lc2=10.77.



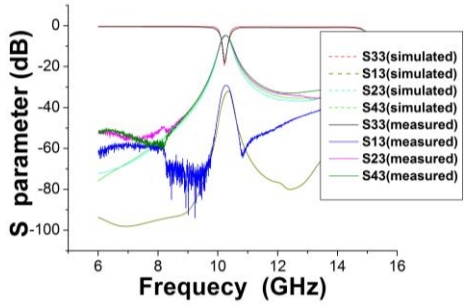
(a)



(b)



(c)



(d)

Fig. 19. Simulated and measured results of the fabricated type B coupler with three SIRCs and a SIFC (a) and (b) Port 1 is excited, (c) and (d) Port 3 is excited.

Figures 19 (a) and (b) shows the simulated and measured results under the in-phase operation. The measured passband is centered at 10.31 GHz with the 1-dB FBW of 2.5%. The minimum insertion losses at the two output ports are (3 + 1.35) and (3 + 1.4) dB, with the

amplitude imbalance of less than 0.1 dB. The return loss is better than 19.3 dB.

Figures 19 (c) and (d) shows the simulated and measured S-parameters under the out-of-phase operation. The measured magnitudes of S23 and S43 are −(3 + 1.4) and −(3 + 1.45) dB, respectively, with the difference of less than 0.1 dB. The measured 1-dB FBW is about 1.9%. The in-band return loss is better than 17.5 dB. As seen, isolation (S31) within the operating frequency band is smaller than −29.6 dB.

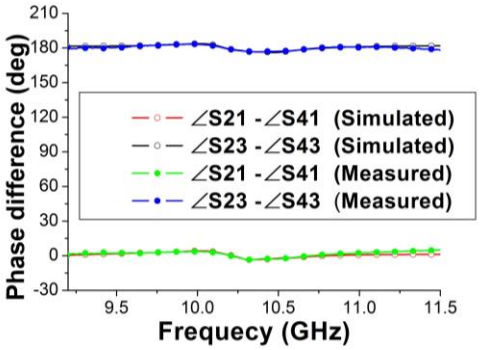


Fig. 20. Phase differences of the fabricated type B coupler with three SIRCs and a SIFC.

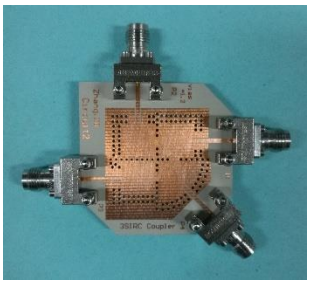


Fig. 21. Photograph of the fabricated filtering type B coupler with three SIRCs and a SIFC.

When signals are injected from port 1 and port 3, the measured in-band phase differences between two output ports are nearly 0° and 180°, respectively, with the variation of less than 3°, as illustrated in Fig.20.

Figure 21 shows the photograph of the fabricated filtering coupler type B based on a SIFC and three SIRCs. A detailed performance comparison with filtering couplers in recent years is shown in Table 1. Compared with [12] and [13], the proposed designs have much higher Q factors and self-consistent electromagnetic shielding structure. By using multilayer structures, the circuits in [20] realize smaller size. However, the multilayer structure is more complex, which increases the difficulty in fabrication and thus increases the processing cost. A filtering rat-race coupler based on the orthogonal modes is proposed in [21]. The desired 0° and 180° phase differences are realized by the inherent characteristics of TE_{102} and TE_{201} modes. Therefore, the input port and isolation port, as well as the two output ports, must be kept perpendicular to each other. Since the coupler performance [21] is limited by port direction, fixed port angle interval makes it unsuitable for other port angular intervals topologies. Furthermore, all ports of the coupler are concentrated in an angle range of 135 degrees, which also leads to EMC problem. In particular, the traditional single type resonators [12-21] structure is not convenient to solve the problem that only part of the ports topology does not match with the desired direction, because it needs to maintain the consistency of the phase for each port. Therefore, the location of components connected to other ports also needs to be adjusted, which will further occupy additional circuit area and affect the overall layout of the system. In this paper, the proposed structures not only have relatively high Q factors, better isolation, simple structure, but also can be suitable for the application of multiple available port angle intervals (APAI), which is helpful to achieve high density and miniaturized RF/microwave wave system.

Table 1: Performance comparison of various filtering rat-race couplers

Ref.	f_0(GHz)/ FBW (%)/ε_r	IL (dB)/ Isolation (dB)/ Q factor	Mag.(dB)/ Phase(deg) Imbalance	APAI *	Circuit Size	Techniques/ Layers*/ Resonators*
[12]	2.4/10/2.2	0.7/20/80	1/2	1	$0.32\times0.32\lambda_g^2$	Microstrip/1/4
[13]	0.47/13/3.38	1.17/25/60	0.2/4.5	1	$0.23\times0.12\lambda_g^2$	Microstrip/1/8
[18]	20/2.6/2.2	1.63/28/190	0.3/5	1	$2.51\times2.51\lambda_g^2$	SIW/1/5
[19]	11/3.6/3.5	1.6/20/170	0.6/8	1	$1.59\times1.26\lambda_g^2$	SIW/1/4
[20]	7.75/2.7/3.5	1.5/25/200	0.6/5	1	$0.79\times0.45\lambda_g^2$	SIW/2/4
[21]	11.8/3.5/3.5	1.3/18/210	0.1/3	1	$1.18\times1.18\lambda_g^2$	SIW/1/3
This work	10.2/3.5/3.5	1.3/27/210	0.05/3	Multiple	$1.19\times1.19\lambda_g^2$	SIW/1/4
	10.28/2.5/3.5	1.38/30/210	0.1/3	Multiple	$1.16\times1.16\lambda_g^2$	SIW/1/4

Where λ_g is the guided wavelength on the substrate at the center frequency f_0, FBW represents the fractional bandwidth. APAI* represents the number of available port angular interval. Layers* represents the number of substrate layers. Resonators* represents the number of resonators.

V. CONCLUSION

In this paper, a novel type of compact band-pass rat-race couplers with flexible port direction based on mixed shape SIW cavities are proposed for the first time. Simulated and measured results have been presented to verify the proposed method. Based on a SIFC and three SIRCs, the other filtering rat-race coupler has been implemented for further verification. Generally, the proposed designs have shown excellent performance of filtering responses, isolation, amplitude balance, 0° and 180° phase differences, as well as the compact structure. In addition, the demand of circuit topology has been firstly integrated into design considerations. With flexible ports configuration but without deterioration of performance, the coupler's topology is easy to match with the desired transmission direction. The proposed filtering couplers with mixed resonators could be more suitable for the development of high density and miniaturized RF/microwave system.

ACKNOWLEDGMENT

This work was supported in part by the Ministry of Science and Technology of the People's Republic of China under Grant 2013YQ200503 and in part by the National Natural Science Foundation of China (NSFC) under Grant 61001028.

REFERENCES

[1] Y. Dong and T. Itoh, "Miniaturized dual-band substrate integrated waveguide filters using complementary split-ring resonators," *IEEE MTT-S Int. Microw. Symp. Dig*., pp. 1-4, June 2011.

[2] X.-P. Chen, K. Wu, and Z.-L. Li, "Dual-band and triple-band substrate integrated waveguide filters with Chebyshev and quasi-elliptic responses," *IEEE Trans. Microw. Theory Techn*., vol. 55, no. 12, pp. 2569-2578, Dec. 2007.

[3] K. Song and Q. Xue, "Novel ultra-wideband (UWB) multilayer slotline power divider with bandpass response," *IEEE Microw. Wirel. Compon. Lett*., vol. 20, no. 1, pp. 13-15, Jan. 2010.

[4] Y. J. Cheng, W. Hong, and K. Wu, "94 GHz substrate integrated monopulse antenna array," *IEEE Trans. Antennas Propag*., vol. 60, no. 1, pp. 121-128, Jan. 2012.

[5] Y. J. Cheng, W. Hong, K. Wu, and Y. Fan, "A hybrid guided-wave structure of half mode substrate integrated waveguide and conductor-backed slotline and its application in directional couplers," *IEEE Microw Wireless Compon Lett*. vol. 21, no. 2, pp. 65-67, Feb. 2011.

[6] Z.-G. Zhang, Y. Fan, Y. J. Cheng, and Y.-H. Zhang, "A compact multilayer dual-mode substrate integrated circular cavity (SICC) filter for X-band application," *Prog. Electromagn. Res*., vol. 122, no. 1, pp. 453-465, Jan. 2012.

[7] Z.-G. Zhang, Y. Fan, and Y.-H. Zhang, "Compact 3-D multilayer substrate integrated circular and elliptic cavities (SICCs and SIECs) dual-mode filter with high selectivity," *Appl. Comp. Electro. Society (ACES) Journal*, vol. 28, no. 4, pp. 333-340, Apr. 2013.

[8] Q. Chen and J. Xu, "Out-of-phase power divider based on two-layer SIW," *Electron Lett*., vol. 50, no. 14, pp. 1005-1007, July 2014.

[9] H. Uchida, N. Yoneda, and S. Makino, "Bandpass directional couplers with electromagnetically-coupled resonators," *IEEE MTT-S Int. Microwave Symp. Digest*. pp. 1563-1566, June 2006.

[10] L.-S. Wu, B. Xia, W.-Y. Yin, and J. Mao, "Collaborative design of a new dual-bandpass 180° hybrid coupler," *IEEE Trans. Microw. Theory Tech*., vol. 61, no. 3, pp. 1053-1066, Mar. 2013.

[11] P. Li, H. Chu, and R. S. Chen, "SIW magic-T with bandpass response," *Electron Lett*., vol. 51, no. 14, pp. 1078-1080, July 2015.

[12] C.-K. Lin and S.-J. Chung, "A compact filtering 180° hybrid," *IEEE Trans. Microw. Theory Tech*., vol. 59, no. 12, pp. 3030-3036, Dec. 2011.

[13] K.-X. Wang, X.-Y. Zhang, S.-Y. Zheng, and Q. Xue, "Compact filtering rat-race hybrid with wide stopband," *IEEE Trans. Microw. Theory Tech*., vol. 63, no. 8, pp. 2250-2560, Aug. 2015.

[14] J.-X. Xu, X.-Y. Zhang, and H.-Y. Li, "Compact narrowband filtering rat-race coupler using quad-mode dielectric resonator," *IEEE Trans. Microw. Theory Tech*., vol. 66, no. 9, pp. 4029-4039, Sep. 2018.

[15] L.-X. Jiao, Y.-L. Wu, Y.-N. Liu, and J.-X. Chen, "Concept for narrow-band filtering rat-race coupler using dual-mode cross-shaped dielectric," *Electron Lett*., vol. 52, no. 3, pp. 212-213, Feb. 2016.

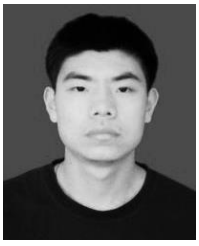
[16] K.-X. Wang, X.-F. Liu, Y.-C. Li, L.-Z. Lin, and X.-L. Zhao, "LTCC filtering rat-race coupler based on eight-line spatially-symmetrical coupled structure," *IEEE Access*, vol. 6, no. 6, pp. 262-269, June 2018.

[17] Z.-G. Zhang, Y. Fan, and Y.-H. Zhang, "Multilayer half-mode substrate integrated waveguide wide-band coupler with high selectivity," *Appl. Comp. Electro. Society (ACES) Journal*, vol. 34, no. 9, pp. 1418-1425, Sep. 2019.

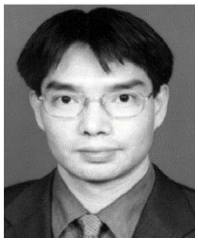
[18] S.-Q. Han, K. Zhou, J.-D. Zhang, C.-X. Zhou, and W. Wu, "Novel substrate integrated waveguide filtering crossover using orthogonal degenerate modes," *IEEE Microw. Wireless Compon. Lett*., vol. 27, no. 9, pp. 803-805, Sep. 2017.

[19] U. Rosenberg, M. Salehi, J. Bornemann, and E. Mehrshahi, "A novel frequency-selective power combiner/divider in single-Layer substrate integrated waveguide technology," *IEEE Microw Wireless Compon Lett*., vol. 23, no. 8, pp. 406-408, Aug. 2013.

[20] Y.-J. Cheng and Y. Fan, "Compact substrate-integrated waveguide bandpass rat-race coupler and its microwave applications," *IET Microw., Antennas Propag*., vol. 6, no. 9, pp. 1000-1006, June 2012.
[21] H.-Y. Li, J.-X. Xu, and X.-Y. Zhang, "Substrate integrated waveguide filtering rat-race Coupler based on orthogonal degenerate modes," *IEEE Trans. Microw. Theory Techn*., vol. 67, no. 1, pp. 140-150, Jan. 2019.
[22] M.-K. Li, C. Chen, and W. Chen, "Miniaturized dual-band filter using dual-capacitively loaded SIW cavities," *IEEE Microw. Wireless Compon. Lett*., vol. 27, no. 4, pp. 344-346, Apr. 2017.
[23] X. Zou, C.-M. Tong, C.-Z. Li, and W.-J. Pang, "Wideband hybrid ring coupler based on half-mode substrate integrated waveguide," *IEEE Microwave and Wireless Components Letters*, vol. 24, no. 9, pp. 596-598, Sep. 2014.
[24] S. Zhang, J.-Y. Rao, J.-S. Hong, and F.-L. Liu, "A novel dual-band controllable bandpass filter based on fan-shaped substrate integrated waveguide," *IEEE Microw. Wireless Compon. Lett*., vol. 28, no. 4, pp. 308-310, Apr. 2018.
[25] Y.-D. Dong and T. Itoh, "Miniaturized substrate integrated waveguide slot antennas based on negative order resonance," *IEEE Trans. Antennas Propag*., vol. 58, no. 12, pp. 3856-3864, Dec. 2010.
[26] M.-C. Tang, S. Q. Xiao, and D. Wang, "Negative index of reflection in planar metamaterial composed of single split-ring resonators," *Appl. Comp. Electro. Society (ACES) Journal*, vol. 26, no. 3, pp. 250-258, Mar. 2011.
[27] M.-C. Tang, H. Wang, L. Guo, X. P. Zeng, H. Liu, and Y. B. Pang, "A compact dual-band patch antenna design based on single-ring split ring resonator," *Appl. Comp. Electro. Society (ACES) Journal*, vol. 31, no. 3, pp. 321-326, Mar. 2016.
[28] Y. S. Li, W. X. Li, and W. H. Yu, "A switchable UWB slot antenna using SIS-HSIR and SIS-SIR for multi-mode wireless communications applications," *Appl. Comp. Electro. Society (ACES) Journal*, vol. 27, no. 4, pp. 340-351, Apr. 2012.
[29] R. Rezaiesarlak, M. Salehi, and E. Mehrshahi, "Hybrid of moment method and mode matching technique for full-wave analysis of SIW circuits," *Appl. Comp. Electro. Society (ACES) Journal*, vol. 26, no. 8, pp. 688-695, Aug. 2011.
[30] Z. J. Zhu, L. Cao, and C. L. Wei, "Novel compact microstrip dual-mode filters with two controllable transmission zeros," *Appl. Comp. Electro. Society (ACES) Journal*, vol. 33, no. 1, pp. 43-48, Jan. 2018.
[31] J.-S. Hong and M.-J. Lancaster, *Microstrip Filter for RF/Microwave Applications*. New York, NY, USA: Wiley; 2001.
[32] R.-J. Cameron, "Advanced coupling matrix synthesis techniques for microwave filters," *IEEE Trans. Microw. Theory Tech*., vol. 51, no. 1, pp. 1-10, Jan. 2003
[33] D.-M. Pozar, *Microwave Engineering*. 2nd edition, New York: Wiley; 1998.

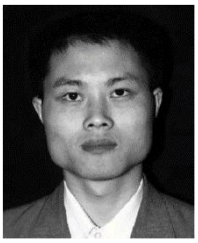


Zhigang Zhang was born in Shanxi Province, China. He received the B.S. degree in Electronic Information Engineering and M.S. degree in Wireless Physics from Sichuan University and is currently working toward the Ph.D. degree in Electro-magnetic Field and Microwave Technology from The University of Electronic Science and Technology of China (UESTC), Chengdu, Sichuan, China. His current research interests include SIW technology and its application, microwave and millimeter-wave filters and couplers, electromagnetic theory.



Yong Fan received the B.E. degree from the Nanjing University of Science and Technology, Nanjing, Jiangsu, China, in 1985, and the M.S. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, Sichuan, China, in 1992.

He fs currently with the School of Electronic Engineering, UESTC. He has authored or coauthored over 60 papers. From 1985 to 1989, he was interested in microwave integrated circuits. Since 1989, his research interests include millimeter-wave communication, electromagnetic theory, millimeter-wave technology, and millimeter-wave systems. Fan is a Senior Member of the Chinese Institute of Electronics (CIE).



Yonghong Zhang received the B.S., M.S., and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1992, 1995, and 2001, respectively. From 1995 to 2002, he was a Teacher with the UESTC. In 2002, he joined the Electronic Engineering Department, Tsinghua University, Beijing, China, as a Doctoral Fellow. In 2004, he rejoined the UESTC. His research interests are in the area of microwave and millimeter-wave technology and applications.

Compact-Size Bandpass Filter for S-Band Transceivers of LEO Satellites

Asmaa E. Farahat and Khalid F. A. Hussein

Microwave Engineering Department
Electronics Research Institute, Cairo, 11843, Egypt
e_asma_e@yahoo.com, Khalid_elgabaly@yahoo.com

Abstract ─ This paper introduces a variety of designs for low-cost and compact-size bandpass filter (BPF) proposed for S-band space-links between ground stations and low earth orbit (LEO) satellites. The proposed designs implement multiple cascaded coplanar waveguide (CPW) resonators with either single or double resonance. The first design is a compact dual-band BPF of dimensions 13×11 mm^2 that can operate simultaneously for both the uplink and downlink in an S-band transceiver. The second design is a single-band BPF of dimensions 25×10 mm^2 that can operate for the space uplink. The third design is a compact single-band BPF of dimensions 17×12 mm^2 that can operate for the space downlink. The three designs are compact size, low cost, and high performance BPFs. The dual-band BPF design is based on dual-resonance mechanism where the resonant structure is constructed as two overlapped quarter-wavelength and half-wavelength CPW resonators; each of them operates in its first-order resonance. The single-band design is based on a quarter-wavelength CPW resonator. Three prototypes are fabricated for the proposed filter designs and their frequency responses are assessed through simulation and experimental measurements showing good agreement. It is shown that the implemented designs satisfy the operational requirements for the assigned space link with high performance, good impedance matching, and low insertion loss.

Index Terms ─ Band-pass filter, CPW resonator, S-band transceiver.

I. INTRODUCTION

S-band communication systems are commonly used in low earth-orbit (LEO) satellites and ground stations for high speed data transfer, for example video or high-quality image data transmission. Usually, an S-band transceiver is used in such type of satellites to provide a high-speed data downlink for payload and low-speed data transfer uplink for telecommand [1], [2]. Due to application requirements, the frequency band for the downlink is much wider and centered at higher frequency than that of the uplink with reasonably wide separation between the two bands. For example, the downlink may be centered at 2250 MHz with bandwidth of 100 MHz, whereas the uplink may be centered at 2050 MHz with a bandwidth of 20 MHz or less. The data rate for the downlink is usually few dozens of Mbps whereas the uplink is few hundreds of Kbps. The data transfer rate and bandwidth are arbitrarily chosen according to the application. In the S-band transmitter of LEO satellites, a BPF is used before the power amplifier to prevent its oscillations and to block the undesired harmonics [1]. In the receiver, to avoid some problems in the frequency mixing stage a BPF is added to eliminate the harmonics and spurious signals or nonlinearities caused by the LNA [2].

Modern S-band communication systems for space links require compact size, high-performance narrow-band bandpass filters (BPF) having low insertion loss, sharp response, and high selectivity together with linear phase or flat group delay in the passband [3]. These filtering features are favorable for the applications that require high spectral efficiency [4-7]. BPF design should also have low radiation loss which can be reduced by using planar resonators of low profile such as the coplanar wave guide (CPW) resonator of narrow slots and narrow central strip. This will enhance the quality factor and reduce the insertion loss [8].

A CPW has the principal advantage that the signal line and the signal grounds are placed on the same substrate surface. This eliminates the need for via holes and, thereby, simplifies the circuit fabrication. Another main advantage of the CPW is that it exhibits lower conductor loss than microstrip lines [9], [10]. CPW parameters are not sensitive to the substrate thickness, and a wide range of impedance is achievable on reasonably thick substrates. Also, circuits can be built using both the odd and the even CPW modes [11]. Moreover, CPWs are open structures, and do not require metallic enclosures [11]. Also, a CPW resonator (CPWR) has its distributed element construction avoiding uncontrolled stray inductances and capacitances, and, thereby, has better microwave properties than a lumped element resonator [12].

Due to their practical importance, microwave and

Submitted On: February 27, 2020
Accepted On: March 26, 2020

millimeter-wave bandpass filter design based on CPWRs has attracted the attention of many research works. In [13], the design of high-quality bandpass filter employing shunt inductively coupled CPW resonators is introduced. In [14], across-coupled CPW structure design is proposed for bandpass filter, which is constructed by cascading several sections of quarter wavelength open-end series stubs. In [15] and [16], compact edge-coupled CPW bandpass filter designs are proposed. In [17], an S-band triangular open loop resonator BPF energized using electromagnetic and edge coupling is presented. In [18], a band pass filter based on folded tri-section stepped impedance resonator (FTSIR) and defected ground structure (DGS) is presented. In [19], a dual-band BPF with several attenuation poles is designed using rectangle structure which exhibits very low insertion loss. In [20], dual-band band-pass filter using tri-section stepped impedance resonators with one stub between parallel coupled lines are employed to realize two pass-bands. In [21], a dual band band-pass filter using multilayer structure is presented.

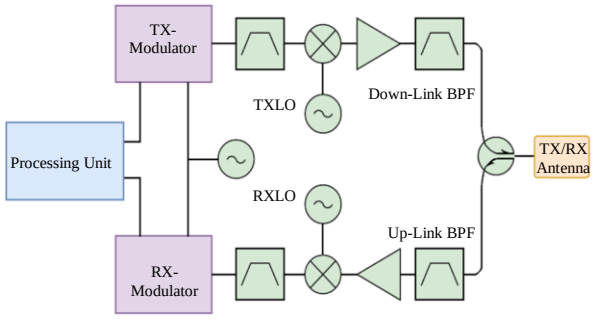
The end-coupled and edge-coupled CPWR structures are commonly used for microwave and millimeter-wave filter design [8]. In end-coupled resonators, the interchange of energy with the coupling gap may be insufficient, even when very narrow gaps are employed. Due to this reason parallel-coupled CPWRs are used in the present work.

The present work proposes two different solutions for narrow-band BPFs required in the S-band communication system of LEO satellites and ground stations for the space uplink and downlink. The first solution is a dual-band BPF design based on a resonant structure constructed as overlapped half-wavelength and quarter-wavelength CPWRs that can work simultaneously for both the uplink and downlink. The other solution provides two designs for a single-band BPF based on quarter-wavelength CPWR; the first design is for a BPF whose passband centered at the operating frequency of the space-uplink, whereas the other design has its passband centered at the operating frequency of the space-downlink.

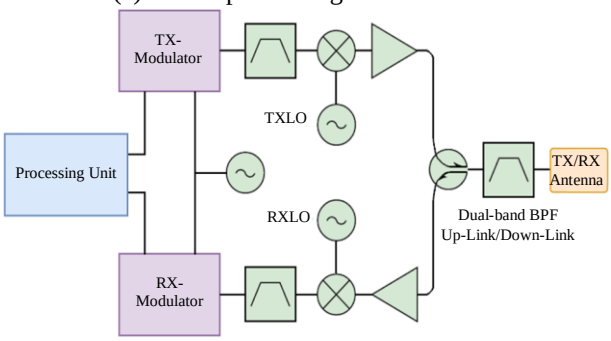
II. PROPOSED BAND-PASS FILTER DESIGNS

The most common functions of the BPF in a transceiver are to limit the bandwidth of operation, to remove the oscillations of the amplifier and to reduce the supurious signals and undesired harmonics. Due to size and weight constraints and power limitations, a small satellite transceiver usually implement the sample antenna for simultaneous transmission and reception. Fig. 1 shows the block diagrams for two possible designs of the sattelite S-band tranceiver. Figure 1 (a) shows a tranceiver system that uses one antenna with two separate BPFs. Figure 1 (b) shows a tranceiver system that employs one antenna with dual-band BPF.

The design of higher-order BPF for each pass-band can be based on using cascaded resonators to mediate coupling between the isolated CPW feeding regions as shown in Fig. 2 (a). Each resonator can be either a single-resonance quarter-wavelength CPWR, or a dual-resonance overlapped half-wavelength and quarter-wavelength CPWRs. In Figs. 2 (b) and 2 (c), single-band and dual-band BPFs constructed using one resonator placed between the two feeding CPW regions are shown, respectively. The design of the single-band and dual-band BPFs are discussed in the following two sections.



(a) Two separate single-band BPFs



(b) One dual-band BPF

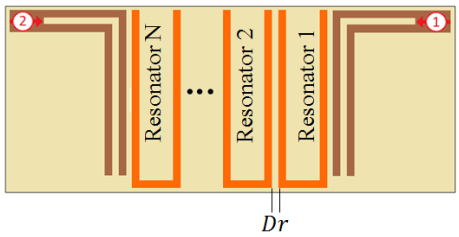
Fig. 1. Block diagrams for two designs of a small satellite S-band transceiver.

According to the above discussion, three designs of a BPF are proposed. The first two designs are single-band BPF, one for the space-uplink, and the other for the space-downlink, whereas the third design is a dual-band BPF that can operate for both uplink and downlink simultaneously.

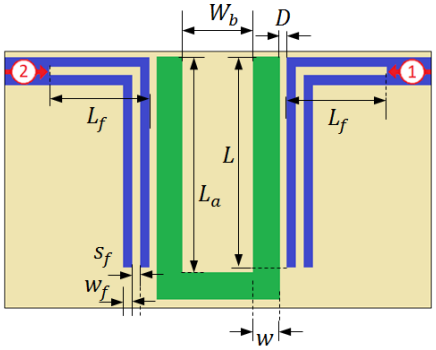
A. Single-band bandpass filters for uplink and downlink

A single-band BPF with the geometry shown in Fig. 2 (b) is proposed for S-band transmitters and receivers of LEO satellites. This filter is constructed as a symmetric structure composed of two feeding (normally isolated) corner-shaped CPW regions which are (capacitively) parallel-coupled to a quarter-wavelength CPWR of

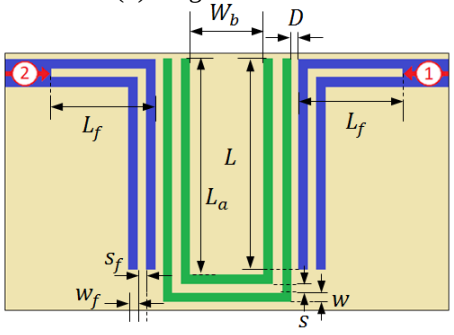
relatively wide central conductor. This CPWR is open-circuited at the base of the U-shape and short-circuited at the other end (top of the U-shape). As long as the operating frequency is far from the resonant frequencies of the CPWR, the two corner-shaped CPW feeding regions are isolated from each other and no power transfer occurs between the filter ports (1) and (2). Only at the resonant frequencies, the CPW feeders are parallel-coupled to the quarter wavelength CPWR and, hence, the power transfer can occur over narrow bands at these frequencies leading to a narrow-band bandpass filter response.



(a) Higher-order BPF based on cascaded CPW resonators



(b) Single-band BPF



(c) Dual-band BPF

Fig. 2. Proposed designs for the S-band BPF for small satellite transceiver

B. Dual-band bandpass filter

A dual-band BPF with the geometry shown in Fig. 2 (c) is proposed for S-band transceivers of LEO satellites. This filter is constructed as two feeding CPW regions parallel-coupled to the dual-resonant CPW structure which is composed of two overlapped resonators. One of them is a half-wavelength CPWR, which is short-circuited at its both ends and forms the perimeter of U-shape. The other resonator is a straight CPWR of relatively wide central conductor formed by the area subtended between the arms of the U-shape. This CPWR is open-circuited at the base of the U-shape and short-circuited at the other end (top of the U-shape) and, thus forms a quarter-wavelength resonator. As long as the operating frequency is far from the resonant frequencies of the dual CPWR structure, the corner-shaped CPW feeding regions are isolated from each other leading to zero power transfer between the filter ports (1) and (2). Power transfer between the filter ports can occur only at the resonant frequencies, when the CPW feeders are parallel-coupled to the dual-resonant CPW structure. As the resonances of the overlapped half-wavelength and quarter-wavelength resonators are shifted from each other, the structure can be dual-band bandpass filter where the frequency separation between the two bands can be controlled by the filter dimensional parameters.

III. THEORETICAL BASICS

A. Half-wavelength and quarter-wavelength CPW resonators

The resonance frequency for a CPWR can be expressed as [22]:

$$f_n^{(1/2)} = \frac{n\,c}{2L_R\sqrt{\varepsilon_{r_{eff}}}}\ ,\ f_n^{(1/4)} = \frac{(2n-1)\,c}{4L_R\sqrt{\varepsilon_{r_{eff}}}}, \tag{1}$$

where L_R is the total length of the resonator, $f_n^{(1/2)}$ and $f_n^{(1/4)}$ are the resonant frequencies of the half-wavelength resonator and quarter-wavelength resonator, respectively, c is the velocity of light in free space, n is the resonance mode order, $n = 1,2,\ldots$, and $\varepsilon_{r_{eff}}$ is the effective dielectric constant of the quasi-TEM mode of the CPW, and L_R is the length of the resonator.

The effective dielectric constant and the characteristic impedance of the quasi-TEM mode of the CPW can be expressed as [22]:

$$\varepsilon_{r_{eff}} = 1 + \frac{\varepsilon_r - 1}{2}\frac{K(\acute{k}_0)}{K(k_0)}\frac{K(k_1)}{K(\acute{k}_1)},$$
$$Z_0 = \frac{30\pi}{\sqrt{\varepsilon_{r_{eff}}}}\frac{K(\acute{k}_0)}{K(k_0)}, \tag{2}$$

where ε_r is the dielectric constant of the substrate material and K denotes the complete elliptic integral of the first kind, which is defined as:

$$K(k) = \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1 - k\sin^2\theta}}. \tag{3}$$

The arguments, k_0, $\acute{k}_0$, k_1, and $\acute{k}_1$, of K are defined as follows:

$$k_0 = \frac{s}{s + 2w}, \tag{4a}$$

$$\acute{k}_0 = \sqrt{1 - k_0^2}, \quad (4b)$$

$$k_1 = \frac{\sinh(\pi s/4h)}{\sinh[\pi(s+2w)/4h]}, \quad (4c)$$

$$\acute{k}_1 = \sqrt{1 - k_1^2}, \quad (4d)$$

where s is the width of the central conductor and w is the width of each side slot.

In spite of being formulated for a CPW of infinitely extending ground, the expressions (2)-(4) can be used as preliminary design rules for a CPW with side ground strips of finite width with good accuracy. According to (2), a 50Ω characteristic impedance of the CPW printed on a substrate of specific material can be obtained by setting the proper values of the strip and slot widths (s, w).

B. Estimation of the quality factor of the bandpass filter

Definitely, the parallel coupling between the resonator and the feeding CPW regions has the effect of external capacitive loading on the resonator. As a consequence, the resonance frequency can be shifted due to such reactive coupling because part of the energy is stored in the electric field of the coupling capacitance. Besides, due to such capacitive coupling, the resultant (loaded) quality factor, Q, is decreased as the coupling to the feeding CPW regions can be considered as a loss channel. The external quality factor Q_e, which is related to the coupling effect as well as the radiation loss dominates the resultant (loaded) quality factor. Thus, the total quality factor can be evaluated through the following relation:

$$\frac{1}{Q} = \frac{1}{Q_u} + \frac{1}{Q_e}, \quad (5)$$

where, Q_u is the self (also, known as internal or unloaded) quality factor of the CPWR. Theoretically, a lossless CPWR has infinite unloaded Q-factor, $Q_u = \infty$. However, practically, Q_u is limited by the conductor and dielectric losses. It should be noted that, for low-loss CPWR, the external Q-factor Q_e, dominates the total Q-factor expressed by (5).

The external Q-factor can be expressed as follows:

$$\frac{1}{Q_e} = \frac{1}{Q_R} + \frac{1}{Q_C}, \quad (6)$$

where Q_R is an equivalent Q-factor related to the radiation loss and Q_C is an equivalent Q-factor related to the reactive coupling between the resonator and feeding CPW regions.

B.1. Calculating the unloaded quality factor of the CPWR

The unloaded Q-factor of both short-circuited and open-circuited half-wavelength CPWR can be expressed as follows [5]:

$$Q_u = \frac{\pi}{2\alpha L_{R_{1/2}}} = \frac{\beta_0}{2\alpha}\sqrt{\varepsilon_{r_{eff}}} = \frac{\omega_0}{2c\alpha}\sqrt{\varepsilon_{r_{eff}}}, \quad (7)$$

where, $L_{R_{1/2}}$ is the length of the half-wavelength resonator, β_0 is the free space wave number, ω_0 is the resonant angular frequency, α is the attenuation constant of the CPW, and $\varepsilon_{r_{eff}}$ is given by (2).

The attenuation constant α of the CPW is related by the conductor and dielectric losses and, hence it can be expressed as follows:

$$\alpha = \alpha_c + \alpha_d, \quad (8)$$

where, α_c is the attenuation caused by the conductor loss whereas α_d is the attenuation caused by the dielectric substrate loss. For a transmission line made of high-conductivity metals like copper ($\sigma = 5.6 \times 10^7 \mathrm{S/m}$) the dielectric loss dominates, which means that $\alpha_d \gg \alpha_c$ and, hence, for a CPW carrying TEM mode, the attenuation constant can be approximated as follows:

$$\alpha \approx \alpha_d = \frac{\omega_0 \tan\delta}{2c}\sqrt{\varepsilon_{r_{eff}}}. \quad (9)$$

Making use of (9), the expression (7) of the unloaded Q-factor of the CPWR reduces to the following:

$$Q_u \approx \frac{1}{\tan\delta}. \quad (10)$$

Substituting from (10) into (5), the total quality factor can be expressed as

$$Q \approx \frac{Q_e}{1 + Q_e \tan\delta}. \quad (11)$$

The last expression can be used to calculate the external Q-factor, Q_e, if the loss tangent of the dielectric substrate is known given that Q has been obtained by simulation.

IV. RESULTS AND DISCUSSIONS

In the present section, both the numerical results obtained by microwave simulation using CST® software package and the experimental results obtained by microwave measurements of some fabricated S-band BPF prototypes are presented, discussed and compared for the purpose of arriving at accurate performance assessment of the proposed filter designs. The first solution proposes a dual-band BPF for both the high-speed payload data downlink and the low-speed telecommand data uplink. The second solution proposes two designs for single-band BPF one for the uplink and the other for the downlink.

A. Comparative performance of multi-resonator BPF

This section is concerned with the performance assessment of a single-band BPF employing single, double, and triple quarter-wavelength resonators arranged as shown in Fig. 2 (a). The design dimensions are $w = 0.4$ mm, $L_a = 15.3$ mm, $L = 14$ mm, $L_f = 4.4$ mm,

$s_f = 0.5$ mm, $w_f = 0.4$ mm, $W_b = 1$mm, $D = 0.2$ mm, and $Dr = 0.1$mm. The frequency responses of the transmission coefficients for three single-band resonators BPF centered at 2.02 GHz are plotted in Fig. 3.

As shown in Fig. 3, it is clear that with increasing the number of cascaded resonators (of the same dimensions), the resulting filter have larger size, wider bandwidth, sharper response and higher rejection ratio in the stop bands. In all the cases, the insertion loss seems to be the same which is low enough to be acceptable for the three designs. However, as it is preferable to have a BPF with compact size and narrow bandwidth for small satellite transceivers as discussed earlier, a single-resonator can be used to implement the required S-band bandpass filter for both single-band and dual-band operations.

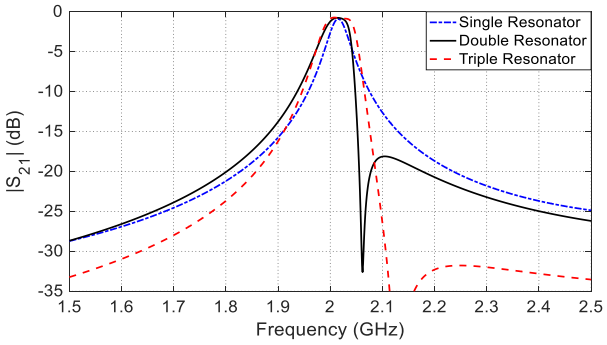


Fig. 3. Frequency response of the transmission coefficient $|S_{21}|$ for single-band BPF designs employing one, two, and three resonators.

B. Dual bandpass filter for S-band transceiver of small satellites

The following presentations and discussions of numerical and experimental results are concerned with a dual-band BPF for S-band transceiver of small LEO satellites as that shown in Fig. 2 (c) designed with the following dimensional parameters, unless otherwise stated. $s = 0.3$ mm, $s_f = 1.1$ mm $w = w_f = 0.4$ mm, $W_b = 2.6$ mm, $L_a = 12.8$ mm, $L_f = 5$ mm, $D = 0.35$ mm, and $L = 12.7$ mm. The substrate material is Rogers RO3010™ with dielectric constant $\varepsilon_r = 10.2$, and height $H = 1.27$ mm. The metal strips and ground are made of copper and have conductivity $\sigma = 5.6 \times 10^7 S/m$.

The frequency response of the transmission coefficient, $|S_{21}|$ and reflection coefficient, $|S_{11}|$ of the proposed dual-band BPF is shown in Fig. 4. The $|S_{21}|$ curve exhibits two peaks. the first peak is at $f = 2.08$ GHz, with frequency bandwidth of 80 MHz, whereas the second peak is at $f = 2.3$ GHz, with frequency bandwidth of 60 MHz. The mechanisms of resonance leading to these values of the resonant frequencies and the corresponding quality factors are explained in the following subsection.

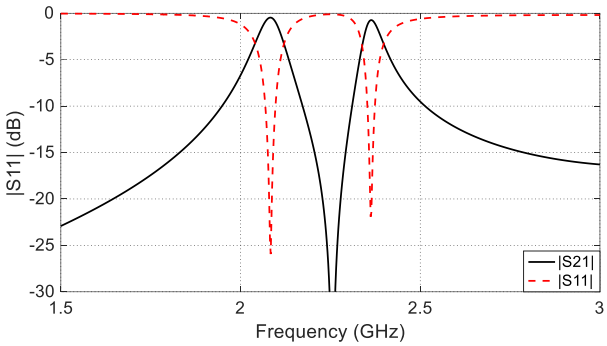


Fig. 4. frequency response of the transmission and reflection coefficients $|S_{21}|$ and $|S_{11}|$ of dual-band BPF designed for S-band tansceiver of a LEO satellite; the lower frequency pass band is for the uplink and the higher frequency pass band is for the downlink.

B.1. Mechanisms of resonance for dual-band operation

The two corner-shaped CPW regions of the main feeder are coupled to each other only at the resonances of each of the two CPWRs formed by the U-shape as mentioned in Section II.B. The frequency response of the filter transfer function has two peaks as shown in Fig. 4. The first peak corresponds to the first resonance of the half-wavelength CPWR on the perimeter of the U-shape which is short-circuit terminated at its both ends. The second peak corresponds to the first resonance of the quarter-wavelength CPWR whose central conductor is the area subtended between the arms of the U-shape. The side (ground) conductors for this CPW are actually the central conductor of the CPW on the perimeter of the U-shape mentioned in section II.B. As shown in Fig. 2 (c), these side conductors have much narrower width in comparison to the width of the central conductor. This resonator is open-ended at the base of the U-shape and is short-circuit terminated at the other end. This is more elaborated showing illustrations of the current and field distributions on the resonator in the following two subsections.

B.1.1. Mechanism of resonance in the lower frequency band

At the lower resonant frequency (2.08 GHz) of the short-ended CPWR formed by the perimeter of the U-shape, the surface current distribution is presented in Fig. 5 (a). It is shown that, like a short-ended half-wavelength resonator, the current has its maximum magnitude at the short-circuited terminals of the resonator whereas the current node of the standing wave is at the middle of the resonator length. The electric field distribution in the slots of this CPW is presented in the same figure which exhibits the behavior of the (even) quasi-TEM mode of the CPW formed by the arms of the U-shape. As this CPWR is half-wavelength whose length is 30.25 mm,

the resonant frequency can be calculated using (1-a), which gives $f_1^{(1/2)} = 2.12$ GHz. The slight deviation of the resonant frequency obtained by simulation from the theoretical value can be attributed to the error due to approximate analytical formula for $\varepsilon_{r_{eff}}$ given by (2) and due to the reactive load caused by coupling to the CPW feeder regions which leads to a shift of the resonant frequency.

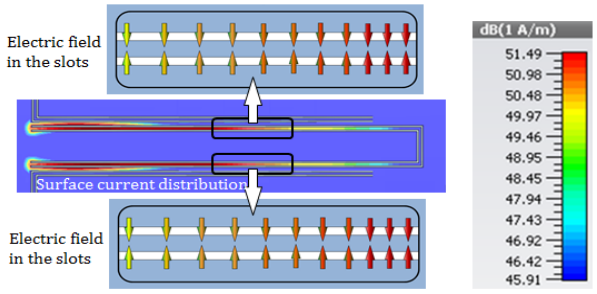
B.1.2. Mechanism of resonance in the higher frequency band

At the higher resonant frequency (2.3 GHz) of the quarter-wavelength CPWR whose central conductor is the region subtended between the arms of the U-shape, the surface current distribution is presented in Fig. 5 (b). This CPW region is open-ended at the base of the U-shape and is short-ended at its other end. It is shown that, like a quarter-wavelength resonator, the current has its maximum magnitude at the short-circuit terminal whereas the current node is at its open-ended. The length of this CPWR is 12.8 mm, and hence, the resonant frequency can be calculated using (1-b), which gives $f_0^{(1/4)} = 2.311$ GHz. The slight deviation of the resonant frequency obtained by simulation from the theoretical value can be attributed to the error of the approximate analytical formula given by (2) for $\varepsilon_{r_{eff}}$ and to the reactive load caused by capacitive coupling of the CPW feeder regions which, in turn, leads to a shift of the resonant frequency.

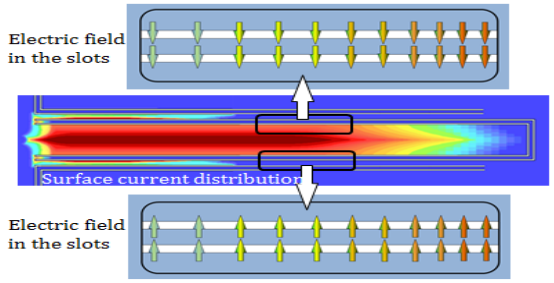
B.2. Experimental assessment of dual-band BPF proposed for S-band transceiver

A prototype of the proposed dual-band BPF is fabricated for experimental investigation of the filter performance. The substrate used for fabrication is Rogers RO3010™, with substrate height h = 1.27 mm, dielectric constant $\varepsilon_r = 10.2$ and dielectric loss tangent $\delta = 0.0021$. The same design dimensions given at the beginning of Section IV.B are used for fabrication. A photograph of the fabricated prototype is presented in Fig. 6 whose size is compared to a coin of the standard one-inch diameter size.

The vector network analyzer (VNA) of Keysight (Agilent) FieldFox N9918A™ is used to measure the transmission and reflection coefficients $|S_{21}|$ and $|S_{11}|$, respectively, of the dual-band BPF prototype under test. For this purpose, the filter prototype is mounted on the substrate test fixture as shown in Fig. 7 (a). After performing the required settings and calibration procedure, the test fixture holding the prototype under test is connected to the VNA as shown in Fig. 7 (b).



(a) Short-ended half-wavelength CPWR at f = 2.08 GHz



(b) Open-ended quarter-wavelength CPW resonator, f = 2.3 GHz

Fig. 5. Surface current on the conductors and electric field in the slots at the resonant frequency

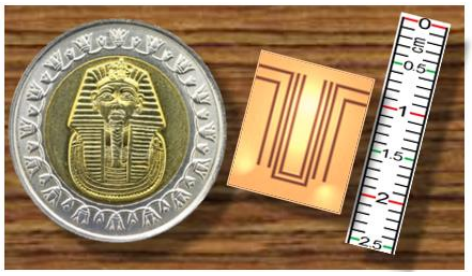
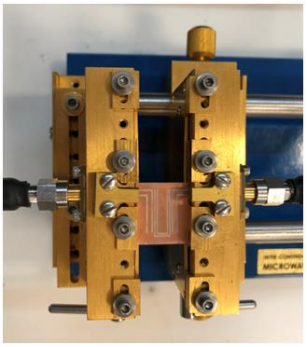
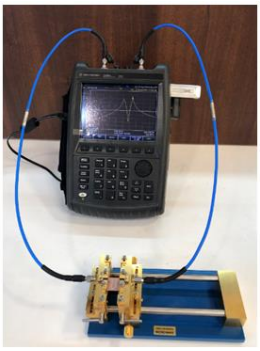


Fig. 6. Photograph of the fabricated prototype of the dual-band BPF proposed for S-band transceivers of LEO satellites.



(a) The fabricated BPF mounted on the VNA test fixture



(b) Measurement of the transmission coefficient $|S_{21}|$ using the VNA

Fig. 7. Measurement of the frequency response of the fabricated prototypes for experimental investigation of the proposed BPF.

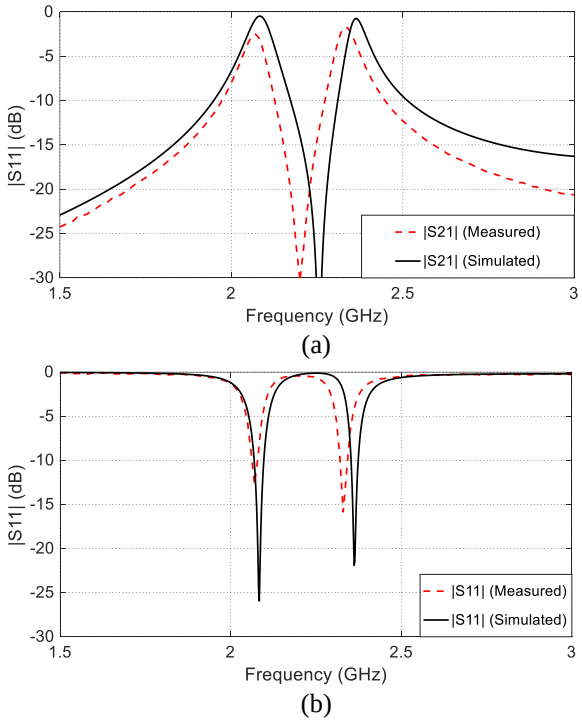


Fig. 8. (a) Frequency response of the transmission coefficient $|S_{21}|$ of the dual-band BPF proposed for S-band transceivers of LEO satellites. (b) Frequency response of the reflection coefficient $|S_{11}|$ of the proposed dual-band BPF for S-band transceivers of LEO satellites.

The frequency response of the transmission coefficient $|S_{21}|$ of the fabricated prototype for the BPF employing half-wavelength CPWR subtending a quarter-wavelength CPWR as measured by the VNA Keysight FieldFox N9918A is presented in Fig. 8 (a) and compared to that obtained by simulation using the CST® software package. Both the experimental and simulation results show good agreement except for a little shift of the resonant frequency, which is, most probably, attributed to the losses encountered in the measurement process. The frequency response of the reflection coefficient $|S_{11}|$ is shown in Fig. 8 (b). The dimensions of the fabricated dual-band BPF are 13×11 mm^2 and hence, it provides a very practical compact and low cost solution for S-band transceivers.

C. Single-band BPF for S-band transceivers of small satellites

Two single-band BPFs having the same design as that shown in Fig. 2 (b), are fabricated with the dimensional parameters listed in Table 1. The single-band BPFs are proposed for S-band transceiver uplink and downlink. The uplink S-band BPF is designed on Rogers RO5880™, with substrate height $h = 1.57$ mm, dielectric constant $\varepsilon_r = 2.2$ and dielectric loss tangent $\delta = 0.0009$. The downlink is designed on Rogers RO3010™, with substrate height $h = 1.27$ mm, dielectric constant $\varepsilon_r = 10.2$ and dielectric loss tangent $\delta = 0.0021$.

Table 1: Design parameters of the proposed single-band BPFs for the S-Band LEO satellite transceiver of small satellites

Dimensional Parameter	Downlink BPF	Uplink BPF
s_f	0.5 mm	1.5 mm
w_f	0.4 mm	0.4 mm
W_b	1.1 mm	1.1 mm
L_a	15 mm	28 mm
L_f	4.05 mm	5.95 mm
D	0.4 mm	0.5 mm
L	13.35 mm	28.3 mm
w	0.4 mm	0.4 mm

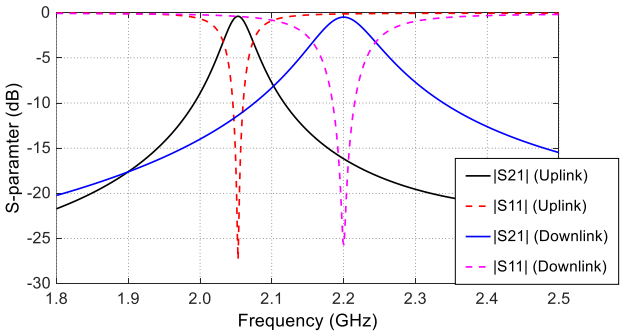


Fig. 9. Frequency responses of the transmission and reflection coefficients of the BPFs proposed for the uplink and the downlink of small satellites S-band transceivers.

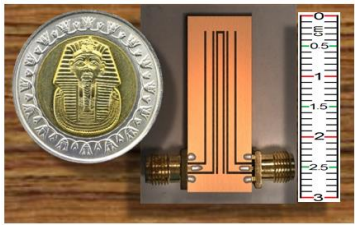
The metal strips and ground are made of copper and have conductivity $\sigma = 5.6 \times 10^7 \mathrm{S/m}$ and metal thickness 0.032 mm. The frequency response of the transmission

coefficients $|S_{21}|$ and the reflection coefficient $|S_{11}|$ of the two BPFs are shown in Fig. 9. The uplink BPF has a center frequency of 2.05 GHz and bandwidth of 43 MHz and the downlink BPF has a 92 MHz bandwidth centered at 2.2 GHz.

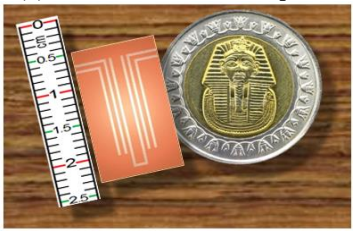
C.1. Experimental assessment of the proposed single-band BPFs

The prototype of the proposed two single band BPFs are fabricated for experimental verifications. The substrate used for fabrication of the single band BPF used in the uplink is Rogers RO5880™, with substrate height h = 1.57 mm, dielectric constant $\varepsilon_r = 2.2$ and dielectric loss tangent $\delta = 0.0009$. The substrate used for fabrication of the single band BPF designed for the downlink is rogers RO3010™, with substrate height h = 1.27 mm, dielectric constant $\varepsilon_r = 10.2$ and dielectric loss tangent $\delta = 0.0021$. The same design dimensions given in table 1 are used in the fabrication process. Photographs of the fabricated prototypes are presented in Fig. 10. The experimental setup is the same as that shown in Fig. 7.

The frequency responses of the transmission and reflection coefficients of the BPF shown in Fig. 10 (a) proposed for the S-band space-uplink are measured by the VNA Keysight FieldFox N9918A and presented in Fig. 11 and compared to those obtained by simulation using the CST® software package. Both the experimental and simulation results show good agreement. The slight differences between the measured and simulated results may be due to losses encountered in the measurement process.



(a) S-band BPF for the uplink



(b) S-band BPF for the downlink

Fig. 10. Fabricated prototype for the proposed single band BPFs used for S-band transceiver.

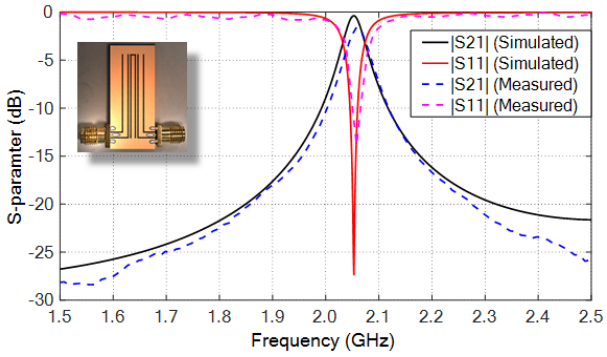


Fig. 11. Frequency responses of the transmission and reflection coefficients $|S_{21}|$ and $|S_{11}|$, for BPF used in the S-band uplink and constructed as two isolated corner-shaped CPW regions partially parallel-coupled to a quarter-wavelength CPWR.

The frequency responses of the transmission and reflection coefficients of the BPF shown in Fig. 10 (b) proposed for the S-band space-downlink are presented in Fig. 12. The center frequency and bandwidth of the proposed single-band BPF can be easily controlled by setting the dimensions of the quarter-wavelength CPWR.

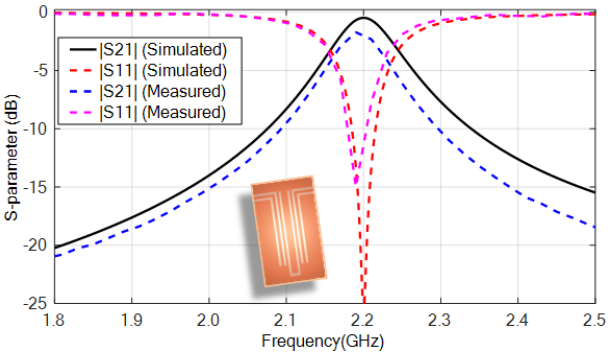


Fig. 12. Frequency responses of the transmission and reflection coefficients $|S_{21}|$ and $|S_{11}|$, respectively, a BPF used in the downlink of S-band tranceiver with the dimensions given in Table 1.

A comparison of the proposed BPF with BPFs found in literature is presented in Table 2.

Table 2: Comparison with similar works

	Size ($\lambda_g \times \lambda_g$)	S_{21}(dB)	S_{11}(dB)
Proposed BPF	0.3 × 0.21	0.63	-27
[17]	0.25 × 0.25	1.3	-25
[23]	0.5 × 0.5	0.1	-25
[24]	0.75 × 0.6	0.04	-23

V. CONCLUSIONS

In this work three compact size, simple structure, low cost, and high performance bandpass filter (BPF) designs for S-band transceiver of LEO satellites are introduced. The first design is a dual-band BPF based on dual-resonance mechanism that can be employed for both the uplink and downlink of the S-band transceiver. The other two designs are based on quarter wavelength CPW resonator. Three prototypes of the proposed filter designs are fabricated and the frequency responses are measured. The experimental measurements show that the implemented designs have high performance with good impedance matching and low insertion loss. The proposed designs are designed on a single layer of the substrate which reduces the substrate losses effect and enhance the quality and performance of the BPF.

The single-band BPF filter is easily controlled as it relies on a single CPWR, its center frequency and quality factor can be accurately tuned. On the other side, the dual band filter is harder to design as the two resonators are overlapped, so it is not easy to tune the center frequency and quality factor for each passband. Employing the dual band BPF or two single band BPFs in an S-band transceiver depends on the transceiver design as mentioned in section 2.

REFERENCES

[1] O. Ceylan, Y. Kurt, F. A. Tunc, H. B. Yagci, and A. R. Aslan, "Low cost S-band communication system design for NANO satellites," *Proceedings of 5th International Conference on Recent Advances in Space Technologies (RAST)*, 2011.

[2] L. Hadj Abderrahmane, M. Benyettou, M. Sweeting, J. R. Cooksley, and P. Garner, "Designing an S-band receiver for LEO applications," *Proceedings of the 11th WSEAS International Conference on Communications*, 26-28, 2007.

[3] J.-S. Hong and M. J. Lancaster, "Couplings of microstrip square open-loop resonators for cross-coupled planar microwave filters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 12, Dec. 1996.

[4] J. S. Hong and M. J. Lancaster, *Microstrip Filters for RF/Microwave Applications*. Wiley, New York, 2001.

[5] D. M. Pozar, *Microwave Engineering*. JohnWiley & Sons, Inc., 2012.

[6] F. Vacondio, A. Ghazisaedi, A. Bononi, and L. A. Rusch, "DQPSK. When is a narrow filter receiver good enough," *IEEE Journal of Lightwave Technology*, vol. 27, no. 22, pp. 5106-5114, Nov. 2009.

[7] N. A. Wahab, N. Hidayat, Z. Ismail Khan, Z. Mat Yassin, N. A. Salim, N. Othman, and M. K. M. Salleh, "Two parallel-coupled rings for narrow bandpass filter application," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 8, no. 373, 2017.

[8] M. Kumar and R. Gowri, "Review on various issues and design topologies of edge coupled coplanar waveguide filters," *Journal of Graphic Era University*, vol. 5, no. 2, pp. 91-96, Jan. 2017.

[9] K. C. Gupta, R. Garg, and I. Bahl, *Microstrip Lines and Slotlines*. Dedham, MA Artech, 1979.

[10] A. Gopinath, "Losses in coplanar waveguides," *IEEE Transactions on Microwave Theory and Techniques*, vol. 30, no. 7, pp. 1101-1104, 1982.

[11] A. V. Sakthivel, A. Sunil, and N. Archak, "Design and implementation of a bandpass coplanar waveguide filter," *Technical Report*, Mar. 2017.

[12] M. Göppl, A. Fragner, M. Baur, R. Bianchetti, S. Filipp, J. M. Fink, P. J. Leek, G. Puebla, L. Steffen, and A. Wallraff, "Coplanar waveguide resonators for circuit quantum electrodynamics," *Journal of Applied Physics*, vol. 104, no. 11, p. 113904, 2008.

[13] J. K. A. Everard and K. K. M. Cheng, "High performance direct coupled bandpass filters on coplanar waveguide," *IEEE Transactions on Microwave Theory and Techniques*, vol. 41, no. 9, pp. 1568-1573, 1993.

[14] R. El Haffar, A. Farkhsi, O. Mahboub, and N. A. Touhami, "Compact size coplanar waveguide bandpass filter design and modeling," *Proceedings of the 2nd International Con.3ference on Computing and Wireless Communication Systems*, (p. 49) ACM, Nov. 2017.

[15] J.-M. Yan, L. Cao, and H.-Y. Zhou, "A novel quad-band bandstop filter based on coupled-line and shorted stub-loaded half-wavelength microstrip resonator," *Progress In Electromagnetics Research*, 65-70, 79, 2018.

[16] O. A. Safia, A. A. Omar, and M. C. Scardelletti, "Design of dual-band bandstop coplanar waveguide filter using uniplanar series-connected resonators," *Progress In Electromagnetics Research*, vol. 27, pp. 93-99, 2011.

[17] B. George, N. S. Bhuvana, and S. K. Menon, "Compact band pass filter using triangular open loop resonator," *Progress In Electromagnetics Research Symposium*, Nov.19-22, 2017.

[18] N. M. Garmjani and N. Komjani, "Improved microstrip olded tri-Section stepped impedance resonator bandpass filter using defected ground structure," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 25, no. 11, Nov. 2010.

[19] L. P. Zhao, X. Zhai, B. Wu, T. Su, W. Xue, and C.-H. Liang, "Novel design of dual-mode bandpass filter using rectangle structure," *Progress In Electromagnetics Research B*, vol. 3, pp. 131-141, 2008.

[20] D. C. Ma, Z. Y. Xiao, L. L. Xiang, X. H. Wu, C. Y.

Huang, and X. Kou, "Compact dual-band bandpass filter using folded SIR with two stubs for WLAN," *Progress In Electromagnetics Research*, vol. 117, pp. 357-364, 2011.

[21] S. Majidifar and S.-V. Makki, "Dual band bandpass filter using multilayer structure," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 30, no. 10, Oct. 2015.

[22] R. N. Simons, "Coplanar waveguide circuits, components, and systems," vol. 165, John Wiley & Sons, 2004.

[23] M. S. S. Subramanian, et al., "Design of dual log-spiral metamaterial resonator for X-band applications," *2012 International Conference on Computing, Communication and Applications (ICCCA)*, 2012.

[24] A. Boutejdar, A. Darwish, and A. Omar, "Design and improvement of compact half-wavelength band pass filter employing overlapped slotted ground structure (SGS) and multilayer technique," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 28, no. 8, Aug. 2013.