APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

April 2025
Vol. 40 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 ARC research group at the Department of Electrical Engineering, Colorado School of Mines

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

APRIL 2025 REVIEWERS

Giovanni Angiulli
Arkaprovo Das
Jogesh Chandra Dash
Sandip Ghosal
Giacomo Giannetti
Shian Hwu
Amir Jafargholi
Ulrich Jakobus
Supakit Kawdungta
Weibin Kong
Giovanni Leone
Leszek Nowosielski
Sangeetha Rangasamy
Natarajamani S.
Ahmed Salman
Ramya Sri
Marsellas Waller
Zhonggen Wang
Erdem Yazgan
Zhihong Ye

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

Vol. 40 No. 4 April 2025

TABLE OF CONTENTS

Fast Direct LDL' Solver for Method of Moments Electric Field Integral Equation Solution

Yoginder Kumar Negi[1], N. Balakrishnan[1], and Sadasiva M. Rao[2]

[1]Supercomputer Education and Research Center
Indian Institute of Science, Bangalore 560012, India
yknegi@gmail.com, balki@iisc.ac.in

[2]Naval Research Laboratory
Washington DC 20375, USA
sadasiva.rao@nrl.navy.mil

Abstract – This paper proposes a new fast direct solver using the block diagonalization method. In our proposed method, the symmetric half single-level compressed block matrix is factorized using the diagonalization method into block diagonal and upper triangle block $\boldsymbol{LDL'}$ format where, due to symmetric property, $\boldsymbol{L}$ is a transpose of $\boldsymbol{L'}$. The far-field blocks in the upper triangle row block are merged and compressed using Adaptive Cross Approximation (ACA) and ***QR*** factorization. The solution consists of solving the diagonal block matrix and matrix-vector multiplication of the compressed row blocks of the upper triangle matrices. Our results show that the factorization cost and memory scales to $O(N^{1.5})$ and the solution process scales to $O(N)$. The method generates an efficient solution process for solving large-scale electromagnetics problems.

Index Terms – Adaptive Cross Approximation (ACA), Electric Field Integral Equation (EFIE), electromagnetics scattering, fast direct solver, matrix compression, Method of Moments (MoM).

I. INTRODUCTION

Method of Moments (MoM) is a well-known integral equation-based Computational Electromagnetics (CEM) [1] method for solving complex electromagnetic radiation and scattering problems numerically in the frequency domain. Compared to differential equation-based methods like Finite Element Method (FEM) [2] and Finite Difference Time Domain (FDTD) [3] methods, MoM is free from grid dispersion error and leads to a smaller matrix size than FEM. The MoM application for solving large-scale electromagnetics problems is limited by dense matrix computation property with the matrix computation and storage cost of $O(N^2)$ for N number of unknowns. Solving the MoM dense matrix with a direct solver leads to $O(N^3)$ computation cost and with iterative solver leads to $N_{itr}\ O(N^2)$ cost for N_{itr} iteration count with $O(N^2)$ matrix-vector product cost. The high matrix computation, storage, and solution cost is mitigated by various matrix compression-based fast solver algorithms proposed by various researchers. The matrix compression for fast solvers may be of two categories: analytical-based compression and algebraic matrix compression. Examples of analytic-based compression methods are Multilevel Fast Multipole Algorithm (MLFMA) [4], Adaptive Integral Method (AIM) [5], and pre-corrected FFT [6]. Similarly, we have algebraic matrix compressed methods like Adaptive Cross Approximation (ACA) [7, 8], IE-QR [9], and H-Matrix [10–12]. Analytical fast solvers are kernel-dependent, whereas algebraic fast solvers are kernel-independent and easy to implement. All the fast solvers work on reducing matrix filling, solution, and storage time. Full wave fast solver in CEM reduces matrix filling and storage time to $O(NlogN)$. However, the matrix solution time depends on the method of solution and whether the algorithm adopts an iterative approach or a direct approach. For the iterative solution, the solution time scales as $N_{itr}\ O(NlogN)$ for $O(NlogN)$ compressed matrix-vector product cost. The solution of these methods relies on the compressed matrix's iterative solution. The iterative solution process is a convergence-dependent method for each matrix-vector product iteration. Furthermore, the convergence for each iteration depends on the condition number of the matrix computed. It is well-known that the Electric Field Integral Equation (EFIE) [13] gives an ill-conditioned MoM matrix. This ill-conditioning due to the closed geometry structure can be overcome by using a Magnetic Field Integral Equation (MFIE) [14] and a Combined Field Integral Equation (CFIE) [15]. Also, the ill-conditioning may be due to structural geometry or mesh quality. The ill-conditioning leads to a high

Submitted On: December 16, 2024
Accepted On: March 26, 2025

https://doi.org/10.13052/2025.ACES.J.400401

iteration count when solved with an iterative solver. The high iteration is mitigated by using various matrix preconditioners [16–18]. The preconditioner computation comes with the extra precondition computation cost and precondition solution cost during the iterative solution process.

Solving multiple right-hand side (RHS) electromagnetics problems like in monostatic radar cross section (RCS) computation or multiport microwave network system analysis with a fast iterative solver may lead to high solution time. The overall solution time will scale to $N_{rhs}\ N_{itr}\ O(NlogN)$ for N_{rhs} RHS. Also, for a multi-RHS or single-RHS system, the number of unknowns increases with the increase in simulation frequency or geometry size. With the increase in the number of unknowns, solving the linear system of equations with an iterative, fast solver will lead to a further increase in the number of iterations (N_{itr}) for the solution process.

The lacuna of the iterative solver for solving linear systems of equations is overcome by using direct matrix solution methods. Direct solvers are the most reliable method for solving any linear system of equations, with a guaranteed solution for the system. However, the high factorization and solution costs hinder the application of direct solvers for significant electromagnetic problems. In the past decade, there has been an inclination among the fast solver research community to develop direct solution-based fast solvers. MLFMA-based analytical fast direct solvers proposed in [19, 20] are kernel-dependent. Algebraic fast direct solvers built upon extended matrix [21] and H^2-Matrix [22] does not scale well for significant problems. A power series-based fast direct solver is proposed in [23, 24], where the solution convergence depends on the matrix's condition number. Sherman-Morrison-Woodbury based fast direct solvers [25, 26] have high factorization and solution time. A ***LU***-based fast direct solver is proposed in [27–29], and factorization is applied to a single-level compressed MoM matrix. ***LU*** matrix factorization and solutions are serial and difficult to parallelize.

This work proposes a new fast direct solver based on the diagonalization applied on a symmetric half single-level compressed block MoM matrix. The factorization cost is reduced by applying low-rank block matrix operation, and the linear cost of the matrix solution is achieved by merging the compressed far-field matrix blocks into a single compressed matrix block. The solution process consists of block diagonal matrix solution and matrix-vector product of the row block compress matrices. The paper is organized as follows: section II presents a brief description of EFIE MoM with block matrix compression. In section III, we present the proposed block diagonalized fast direct solver. Section III also presents the low-rank matrix operation performed on the compressed matrices to reduce the factorization and solution time. In section IV, the efficiency and accuracy of the proposed fast direct solver is presented. Section V concludes the paper.

II. BRIEF DESCRIPTION OF EFIE-MOM

The MoM matrix is computed for 3D surfaces using EFIE, MFIE, or CFIE formulations. Selection of the integral equation method is essential when solving the matrix with a regular iterative solver. MFIE is only applicable for closed-body geometries. CFIE is a combination of EFIE and MFIE, which further increases matrix computation time. For the sake of clarity, this work uses only EFIE for MoM matrix computation to solve 3D Perfect Electric Conductor (PEC) geometry and can be easily extended to MFIE and CFIE. The governing equation for EFIE is:

$$E_{total} = E_{inc} + E_{scatt}\ , \tag{1}$$

where E_{total} is the total electric field equal to the sum of the incident electric field (E_{inc}) and the scattered electric field (E_{scatt}). Applying the boundary condition, expanding current density ($J(r')$) and charge density ($\rho(r')$) for the electric field vector potential and scalar potential with the RWG [30] basis functions (f_i), and performing Galerkin testing, the elements of the MOM matrix are:

$$\boldsymbol{Z}(i,j) = \frac{j\omega\mu}{4\pi}\iint f_i \cdot f_j \frac{e^{-jk|r-r'|}}{r-r'} dsdt + \frac{1}{j\omega 4\pi\varepsilon}\iint \nabla \cdot f_i \frac{e^{-jk|r-r'|}}{r-r'} \nabla \cdot f_j dsdt. \tag{2}$$

In the above equation, the MoM matrix element is computed for the i^{th} test j^{th} source basis. In equation (2), k is the wavenumber, r and r' represent the observer and source points, and μ and ε are the permeability and permittivity of the background material. Integration is performed over the RWG source and testing domains. MoM matrix computation using equation (2) leads to a linear system of equations. The system of equations can be written as a combination of a near-field interaction matrix $\boldsymbol{Z_N}$ and a far-field interaction matrix $\boldsymbol{Z_F}$ given by:

$$[\boldsymbol{Z_N} + \boldsymbol{Z_F}\ ]\boldsymbol{x} = \boldsymbol{b}. \tag{3}$$

Solving equation (3) for $\boldsymbol{x}$ presents computation time and memory limitations, which can be overcome by applying various fast solver methods. These methods work on the compressibility of the far-field matrix.

For the computation of a fast solver, the mesh of the 3D geometry is divided into small blocks using the multilevel binary-tree or oct-tree partition method [12]. The fast solver is the single level when the far-field matrix compression is carried out at the lowest level, and the multilevel is when the matrix compression is done at all levels. Single-level fast solver reduces matrix filling and solution time to $O(N^{1.5})$ whereas multilevel reduces the

time complexity to $O(NlogN)$. This work uses a single-level compressed MoM block matrix fast solver to ease block matrix operation. For the single-level fast solver matrix construction, matrix compression is applied at the lowest level of binary-tree-based 3D geometry decomposition, and the interaction is computed for the mesh block satisfying the admissibility condition [12]. At the leaf level, the block interaction that does not satisfy the admissibility condition is considered a near-field interaction. Further, this work employs the re-compressed ACA method [31] to compute a symmetric single-level compressed block matrix. The compressed matrix is made symmetric by averaging the upper diagonal and lower diagonal values and replacing the upper diagonal and lower diagonal values with the average value. Using the symmetric property, we can reuse the upper diagonal value for matrix-vector product and matrix factorization. Exploiting the symmetric property, we compute and save the diagonal and upper diagonal matrix in the compressed far-field and dense near-field blocks [12]. The single-level symmetric half-compressed block matrix is used for factorization and solution process. The formulation for the new fast direct solver with reduced factorization and solution time is discussed in the next section.

III. FAST DIRECT SOLVER

This section discusses the factorization and solution process for developing a fast direct solver applied on a single-level compressed block matrix. The matrix diagonalization process has previously been used to solve the linear system of equations for sparse and dense matrices [32–34]. However, these methods lead to a high cost of factorization and solution. In our previous work [18, 23, 24], the diagonalization of the near-field block matrix is discussed in detail. There, the computation is performed on the dense near-field block matrices. Due to the dense nature of the near-field matrix, the cost of computation and storage is kept low. Extending the diagonalization method to a fast solver-based full matrix is limited by the far-field matrix blocks. In this work, we diagonalize the full MoM compressed matrix. To understand the complete implementation of the factorization process, we will discuss the factorization process for the whole matrix in detail. The far-field matrix operation cost is reduced by performing low-rank block matrix operation. The method is applied to the MoM matrix to factorize it to $\boldsymbol{LDL'}$ format, where $\boldsymbol{D}$ is a diagonal block matrix and $\boldsymbol{L}$ and $\boldsymbol{L'}$ are lower and upper triangle block matrices. The lower triangle block matrix $\boldsymbol{L}$ is the transpose of the upper triangle block matrix, leading to computation and memory savings. In the following subsection, details of the diagonalization process, along with the ways to make it faster for set-up and solution matrix-vector product, are presented.

A. Block matrix factorization

Gaussian Elimination is a well-known method for diagonalizing dense or sparse block matrices. The Gaussian Elimination is performed over the single-level symmetric compressed block matrix. An asymmetric compressed block matrix is computed for the mesh geometry divided into mesh element clusters based on the multilevel binary-tree division. The single-level compressed block matrix near-field and far-field matrix interaction is computed for binary-tree mesh blocks at the lowest level. The symmetric half-block matrix $[\boldsymbol{Z}]$ for factorization is shown below:

$$[\boldsymbol{Z}] = \begin{bmatrix} \boldsymbol{Z}_{11} & \boldsymbol{Z}_{12} & \boldsymbol{Z}_{13} & \boldsymbol{Z}_{14} \\ 0 & \boldsymbol{Z}_{22} & \boldsymbol{Z}_{23} & \boldsymbol{Z}_{24} \\ 0 & 0 & \boldsymbol{Z}_{33} & \boldsymbol{Z}_{34} \\ 0 & 0 & 0 & \boldsymbol{Z}_{44} \end{bmatrix}. \tag{4}$$

The above block matrix consists of near- and far-field matrix blocks. The factorization process consists of diagonalizing the above matrix by multiplying it with the right sparse block matrix. The right scaling matrix nullifies the row blocks of the matrix, leaving a diagonal block matrix. Right scaling matrix $[\boldsymbol{\alpha_1}]$ for first row blocks is given as:

$$[\boldsymbol{\alpha}_1] = \begin{bmatrix} \boldsymbol{I}_{11} & -\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{12} & -\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{13} & -\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{14} \\ 0 & \boldsymbol{I}_{22} & 0 & 0 \\ 0 & 0 & \boldsymbol{I}_{33} & 0 \\ 0 & 0 & 0 & \boldsymbol{I}_{44} \end{bmatrix}, \tag{5}$$

where $\boldsymbol{I}_{11}$, $\boldsymbol{I}_{22}$, $\boldsymbol{I}_{33}$, and $\boldsymbol{I}_{44}$ are the identity block matrices. Equations (4) and (5) are combined to scale the first-row blocks of $[\boldsymbol{Z}]$ to diagonal blocks, and the system of the equation is given as:

$$[\widetilde{\boldsymbol{Z}}^1] = \begin{bmatrix} \boldsymbol{Z}_{11} & \boldsymbol{Z}_{12} & \boldsymbol{Z}_{13} & \boldsymbol{Z}_{14} \\ 0 & \boldsymbol{Z}_{22} & \boldsymbol{Z}_{23} & \boldsymbol{Z}_{24} \\ 0 & 0 & \boldsymbol{Z}_{33} & \boldsymbol{Z}_{34} \\ 0 & 0 & 0 & \boldsymbol{Z}_{44} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{I}_{11} & -\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{12} & -\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{13} & -\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{14} \\ 0 & \boldsymbol{I}_{22} & 0 & 0 \\ 0 & 0 & \boldsymbol{I}_{33} & 0 \\ 0 & 0 & 0 & \boldsymbol{I}_{44} \end{bmatrix}. \tag{6}$$

Equation (6) is represented as:

$$[\widetilde{\boldsymbol{Z}}^1] = [\boldsymbol{Z}]\,[\boldsymbol{\alpha}_1]. \tag{7}$$

Performing the block matrix and scaling matrix multiplication in the above equation, the first-row block diagonalized matrix block equation is written as:

$$\left[\widetilde{\boldsymbol{Z}}^1\right] = \begin{bmatrix} \boldsymbol{Z}_{11} & 0 & 0 & 0 \\ 0 & \boldsymbol{Z}_{22}-\boldsymbol{Z}_{21}\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{12} & \boldsymbol{Z}_{23}-\boldsymbol{Z}_{21}\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{13} & \boldsymbol{Z}_{24}-\boldsymbol{Z}_{21}\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{14} \\ 0 & 0 & \boldsymbol{Z}_{33}-\boldsymbol{Z}_{31}\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{13} & \boldsymbol{Z}_{34}-\boldsymbol{Z}_{31}\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{14} \\ 0 & 0 & 0 & \boldsymbol{Z}_{44}-\boldsymbol{Z}_{41}\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{14} \end{bmatrix}. \tag{8}$$

Equation (8) gives the diagonalization of the first-row blocks. In the above equation, using the symmetry property, the matrix blocks $\boldsymbol{Z}_{21}$, $\boldsymbol{Z}_{31}$, and $\boldsymbol{Z}_{41}$ are obtained by computing the transpose of $\boldsymbol{Z}_{12}$, $\boldsymbol{Z}_{13}$, and $\boldsymbol{Z}_{14}$. So $\boldsymbol{Z}_{21} = \boldsymbol{Z}'_{12}$, $\boldsymbol{Z}_{31} = \boldsymbol{Z}'_{13}$, and $\boldsymbol{Z}_{41} = \boldsymbol{Z}'_{14}$. Likewise, each row block is diagonalized by computing the row scaling matrix block and multiplying it with the row diagonalized matrix. The final diagonalization process is given as:

$$[\boldsymbol{Z_D}] = [\boldsymbol{\alpha}_3^{\mathrm{T}}][\boldsymbol{\alpha}_2^{\mathrm{T}}][\boldsymbol{\alpha}_1^{\mathrm{T}}][\boldsymbol{Z}][\boldsymbol{\alpha}_1][\boldsymbol{\alpha}_2]][\boldsymbol{\alpha}_3] \,. \tag{9}$$

Equation (9) written in diagonal form is:

$$[\boldsymbol{Z_D}] = \begin{bmatrix} \boldsymbol{Z}_{11} & 0 & 0 & 0 \\ 0 & \widetilde{\boldsymbol{Z}}_{22} & 0 & 0 \\ 0 & 0 & \widetilde{\boldsymbol{Z}}_{33} & 0 \\ 0 & 0 & 0 & \widetilde{\boldsymbol{Z}}_{44} \end{bmatrix}. \tag{10}$$

In equation (10), $[\boldsymbol{Z_D}]$ is the diagonal block matrix $\boldsymbol{D}$ and $[\boldsymbol{\alpha}_1][\boldsymbol{\alpha}_2]][\boldsymbol{\alpha}_3]$ is the upper triangle block matrix $\boldsymbol{U}$ of the $\boldsymbol{LDU}$ factorization. $\boldsymbol{L}$ is the transpose of $\boldsymbol{U}$ due to the symmetric property of the matrix. The block diagonalized system of the equation is written as:

$$[\boldsymbol{Z_D}]\,[\widetilde{\boldsymbol{x}}] = \left[\widetilde{\boldsymbol{b}}\right]. \tag{11}$$

Here, $[\boldsymbol{b}]$ and $[\boldsymbol{x}]$ is computed by:

$$\left[\widetilde{\boldsymbol{b}}\right] = [\boldsymbol{\alpha}_3^{\mathrm{T}}][\boldsymbol{\alpha}_2^{\mathrm{T}}][\boldsymbol{\alpha}_1^{\mathrm{T}}][\boldsymbol{b}], \tag{12}$$

$$[\boldsymbol{x}]P = [\boldsymbol{\alpha}_1][\boldsymbol{\alpha}_2][\boldsymbol{\alpha}_3][\widetilde{\boldsymbol{x}}]. \tag{13}$$

The final solution process consists of solving equation (11) and performing block matrix-vector products to extract the solution vector in equations (12) and (13). The matrix solution cost reduction will be discussed later.

The major cost of the above-discussed process is the factorization process. The process includes the computation of the block scaling matrix in equation (5) and performing block matrix operations given in equation (8). The generalized matrix operation in equation (8) for scaling k^{th} row block operating on varying m^{th} row and n^{th} column matrix blocks is written as:

$$\widetilde{\boldsymbol{Z}}_{mn} = \boldsymbol{Z}_{mn} - \boldsymbol{Z}_{mk}\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}. \tag{14}$$

In equation (14), for nullifying one row, there are four major matrix operations to be performed on the remaining matrix blocks. The matrix operations are: matrix inversion of the diagonal block matrix $[\boldsymbol{Z}_{kk}^{-1}]$, matrix solution for block matrix $[\boldsymbol{Z}_{kn}]$, matrix multiplication, and block matrix addition. To perform the above operations in equation (8), matrix inversion needs to be computed once; a matrix block solution is required for the selected row block to be nullified. The major matrix operations to be repeated for each block computation are matrix multiplication and addition. When all the blocks are dense, the operations are straightforward, but this will lead to a high factorization cost, along with a high matrix computation cost. In this work, we reduced the cost of matrix computation and solution as follows.

The matrix consists of dense near-field matrix blocks and low-rank compressed far-field matrix blocks, as given in equation (3). The near-field dense matrix operations are performed using equation (14). However, to keep the computation cost low, we will perform a low-rank block matrix operation [35] on the far-field compressed form matrix. From equation (14), the low-rank matrix operating will be required for block matrix solution for the computation of scaling matrix, block matrix multiplication, and block matrix addition. The low-rank matrix operations are discussed in further subsections.

1. Low-rank matrix solution

The low-rank matrix solution is computed in equation (14) for m^{th} row and n^{th} column block as $[\boldsymbol{Z}_{kk}^{-1}][\boldsymbol{Z}_{kn}]$, where $[\boldsymbol{Z}_{kk}^{-1}]$ is $\boldsymbol{LU}$ factorized dense matrix and $[\boldsymbol{Z}_{kn}]$ is a low-rank compressed matrix of the form $[\boldsymbol{U}]_{kn}[\boldsymbol{V}]_{kn}$, and $[\boldsymbol{U}]_{kn}$ is of size $u \times r$ and $[\boldsymbol{V}]_{kn}$ is of size $r \times v$, for u row size, v column size, and r rank of the block matrix such that $r'\ min(m,n)$. Solving $[\boldsymbol{Z}_{kk}^{-1}]$ for $[\boldsymbol{U}]_{kn}$ and multiplying with $[\boldsymbol{V}]_{kn}$ is equal to solving $[\boldsymbol{Z}_{kk}^{-1}]$ for $[\boldsymbol{Z}_{kn}]$ as shown in Fig. 1. The process saves the solution time by avoiding a full matrix solution.

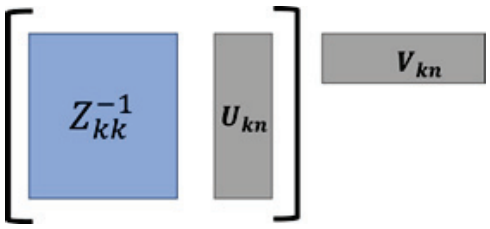


Fig. 1. Low-rank block matrix solution. The matrix operation inside the brackets is performed first to reduce the computation cost.

2. Low-rank matrix multiplication

Low-rank matrix multiplication is computed in equation (14) for m^{th} row and n^{th}column block for $[\boldsymbol{Z}_{mk}]$ and $[\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}]$. The low-rank multiplication in this operation comes with two scenarios. First, $[\boldsymbol{Z}_{mk}]$ is dense and $[\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}]$ is in compressed form and, second, both $[\boldsymbol{Z}_{mk}]$ and $[\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}]$ are in compressed form.

In the first case, let us consider $[\boldsymbol{Z}_{mk}]$ of size $u2 \times u1$ and $[\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}]$ is in the compressed form of $[\boldsymbol{U}_1]_{u1\times r1}[\boldsymbol{V}_1]_{r1\times v1}$ where $u1$ is row size, $v1$ is column size, and $r1$ is the rank of the block matrix. Compressed fast multiplication is carried out by multiplying $[\boldsymbol{Z}_{mk}]$ with $[\boldsymbol{U}_1]$ and replacing it with $[\boldsymbol{U}_1]$. Multiplication cost is reduced to $u2 \times u1 \times r1$. Similarly, compressed matrix multiplication is performed when $[\boldsymbol{Z}_{mk}]$ is in compressed form and $[\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}]$ is in dense form.

In the second case, when $[\boldsymbol{Z}_{mk}]$ and $[\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}]$ both are in compressed form, fast matrix multiplication is

carried out by multiplying compressed parts. Let $[\boldsymbol{Z}_{mk}]$ be the compressed form of $[\boldsymbol{U}_1]_{u1\times r1}[\boldsymbol{V}_1]_{r1\times v1}$ where $u1$ is row size, $v1$ is the column size, and $r1$ is the rank of the matrix. Let $[\boldsymbol{Z}_{mk}]$ be the compressed form of $[\boldsymbol{U}_2]_{u2\times r2}[\boldsymbol{V}_2]_{r2\times v2}$ where $u2$ is row size, $v2$ is column size, and $r2$ is the rank of the matrix block. Fast matrix multiplication is carried out by multiplying $[\boldsymbol{V}_1]_{r1\times v1}$ with $[\boldsymbol{U}_2]_{u2\times r2}$ as shown in Fig. 2, leading to a block matrix of the size $r1 \times r2$. The low-rank matrix is multiplied by row block $[\boldsymbol{U}_1]_{u1\times r1}$ or column block $[\boldsymbol{V}_2]_{r2\times v2}$ and replacing the row or column block.

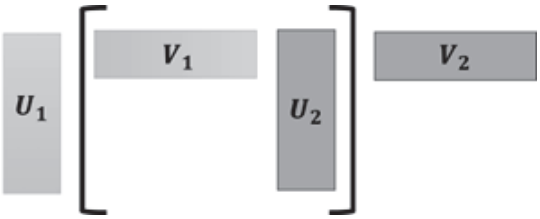


Fig. 2. Low-rank block matrix multiplication operation. Multiplication inside the brackets is performed first for cost saving.

3. Low-rank matrix addition

Low-rank block matrix addition computation for m^{th} row is required for adding matrix blocks $[\boldsymbol{Z}_{mn}]$ and $-[\boldsymbol{Z}_{mk}\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}]$ in equation (14). Converting the compressed matrices in dense form and adding them will lead to high cost, and adding only the respective compressed row and column blocks leads to an incorrect reluctant matrix. Low-rank matrix operation is performed by merging the respective column row and column blocks and compressing them. Let $[\boldsymbol{Z}_{mn}]$ be of the form $[\boldsymbol{U}_1]_{u1\times r1}[\boldsymbol{V}_1]_{r1\times v1}$ and $-[\boldsymbol{Z}_{mk}\boldsymbol{Z}_{kk}^{-1}\boldsymbol{Z}_{kn}]$ of the form $[\boldsymbol{U}_2]_{u1\times r2}[\boldsymbol{V}_2]_{r2\times v1}$. Then $[\boldsymbol{U}_1]$ and $[\boldsymbol{U}_2]$ is merged by column, and $[\boldsymbol{V}_1]$ and $[\boldsymbol{V}_2]$ is merged by row as shown in Fig. 3. The merged matrices are low-rank matrices and are further compressed and converted into row and column blocks by following the matrix multiplication operation discussed in the previous step.

The low-rank matrix operation discussed above reduces the factorization cost for the full single-level compressed matrix. Factorization is carried out with a row block-wise operation by replacing the existing compressed and dense matrices. This process leads to memory savings with no extra storage requirements for matrix storage. The single-level compressed and dense matrix is replaced by a factorized compressed and dense matrix. The solution time for the factorization can be further reduced by merging the compressed factorized far-field matrix blocks, which are discussed later.

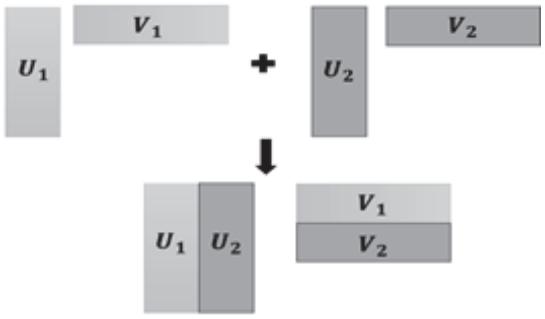


Fig. 3. Low-rank block matrix row and column merging are used to perform low-rank block matrix addition operations.

B. Fast matrix solution

The solution process includes solving the diagonal block matrix in equation (11) and the matrix-vector product in equations (12) and (13). The solution cost of the diagonal block is $O(N)$, as shown in [18], where it is used as a preconditioner. The scaling row block matrices $[\boldsymbol{\alpha}_1]$, $[\boldsymbol{\alpha}_1]$, and $[\boldsymbol{\alpha}_3]$ consist of dense near-field and compressed far-field blocks and for the compressed far-field and dense near-field matrix-vector product cost scales to $O(N^{1.5})$. The matrix-vector product cost is further reduced by merging the far-field compressed blocks as done in H^2-Matrix computation. In the scaling matrix block equation (6), let us suppose all matrix blocks $[-\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{12}]$, $[-\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{13}]$, and $[-\boldsymbol{Z}_{11}^{-1}\boldsymbol{Z}_{14}]$ are in compressed form and are represented as $[\boldsymbol{U}_1]_{u1\times r1}[\boldsymbol{V}_1]_{r1\times v1}$, $[\boldsymbol{U}_2]_{u2\times r2}[\boldsymbol{V}_2]_{r2\times v2}$, and $[\boldsymbol{U}_3]_{u3\times r3}[\boldsymbol{V}_3]_{r3\times v3}$ of the same row basis and size such that $u1 = u2 = u3$. For merging, we follow the following steps:

1. Compute $\boldsymbol{QR}$ decomposition of the first column block matrix. $[\boldsymbol{U}_1]_{u1\times r1}$ into $[\boldsymbol{Q}_1]_{u1\times r1}[\boldsymbol{R}_1]_{r1\times r1}$ where $\boldsymbol{Q}_1$ is an orthonormal matrix and $\boldsymbol{R}_1$ is an upper triangle block matrix.
2. The first compressed row block matrix $[\boldsymbol{V}_1]_{r1\times v1}$ is replaced with the new matrix product $[\boldsymbol{R}_1] \times [\boldsymbol{V}_1]$ represented as $\left[\widetilde{\boldsymbol{V}}_1\right]_{r1\times v1}$.
3. The second matrix block $[\boldsymbol{V}_2]_{r2\times v2}$ is scaled to $[\boldsymbol{Q}_1]$ using transfer matrix computed as $[\boldsymbol{Q}_1]^{-1}\ [\boldsymbol{U}_2]$. As $[\boldsymbol{Q}_1]$ is an orthonormal matrix, the solution process will be $[\boldsymbol{Q}_1]^{'}\ [\boldsymbol{U}_2]$. The transfer matrix will be of size $r1 \times r2$ and is multiplied with $[\boldsymbol{V}_2]_{r2\times v2}$ represented as $\left[\widetilde{\boldsymbol{V}}_2\right]_{r1\times v2}$.

Following the above steps, the third block compressed row matrix $[\boldsymbol{V}_3]_{r3\times v3}$ is scaled to the first block using a transfer matrix and is represented as $\left[\widetilde{\boldsymbol{V}}_3\right]_{r1\times v3}$.The conversion process is pictorially represented in Fig. 4.

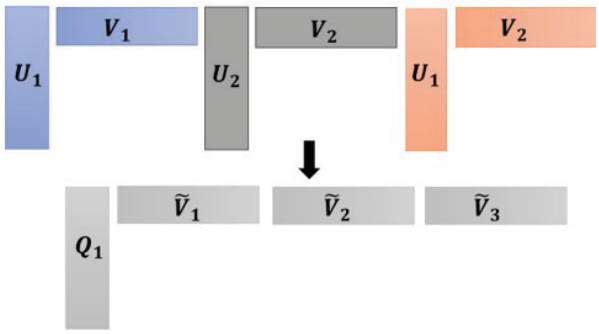


Fig. 4. Conversion of multiple compressed far-field blocks to a single compressed block to reduce matrix-vector product cost.

The new scaling matrix requires storage of the column block matrix. $[\boldsymbol{Q}_1]$ and row block matrices $\left[\widetilde{\boldsymbol{V}}_1\right]$, $\left[\widetilde{\boldsymbol{V}}_2\right]$, and $\left[\widetilde{\boldsymbol{V}}_3\right]$. The row block matrix is merged to form a new matrix $\left[\widetilde{\boldsymbol{V}}_{\text{new}}\right]$. The final compressed matrix will be of the form $[\boldsymbol{Q}_1] \times \left[\widetilde{\boldsymbol{V}}_{new}\right]$. This reduces the overall matrix storage and matrix-vector product cost. Our results show the matrix-vector product cost scaled to $O(N)$.

C. Complexity analysis

Complexity analysis for the near field factorization is discussed in [18]. In this section we will add the far-field operation cost on the proposed method. The proposed algorithm has two steps. First is the low rank compressed single level matrix factorization using low rank matrix operation and second is low rank matrix merging and solution using matrix-vector product. Let the low rank matrix be of size of rank k with row and column size n such that $k << n$.

Step 1. The factorization part consists of low rank matrix solutions. The solution cost for the low rank matrix will be reduced to kn, as the solution is carried out for compressed column matrix. The solution is followed by low rank matrix multiplication operation, where the matrix is merged to form a matrix of size k^2n which is further compressed with complexity of $k^2(n+n)$, retaining the low cost of operation.

Step 2. The merging part includes transpose of the low rank row block and multiplication. The transpose is of $O(1)$ complexity whereas multiplication cost is k^2n. The solution part is of low rank matrix-vector product costing $k^2\,(n+n)$.

Our experimental results for a hollow cylinder of radius 0.25λ with varying lengths and unknowns show that the factorization cost remains to be $O(N^{1.5})$ complexity and total solution time costs maintain $O(N)$ complexity. The complexity is shown in Fig. 5 for an increasing number of unknowns. The cylinder is meshed with tringles for $\lambda/10$ edge length. The binary tree division for mesh is terminated for a block size of 10λ-15λ at the lowest level.

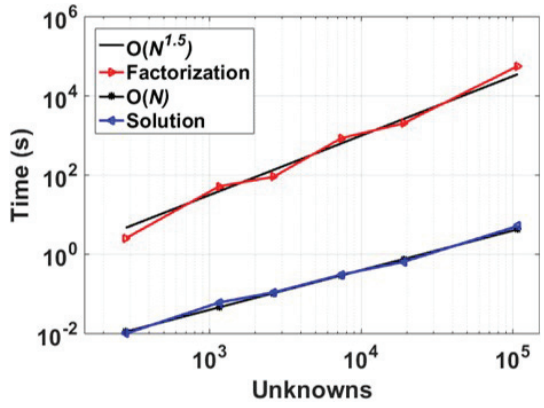


Fig. 5. Block matrix factorization and upper triangle block time using an increasing number of unknowns.

Figure 6 shows memory complexity for factorization process with an increasing number of unknowns. It is observed that the process gives $O(N^{1.5})$ memory complexity.

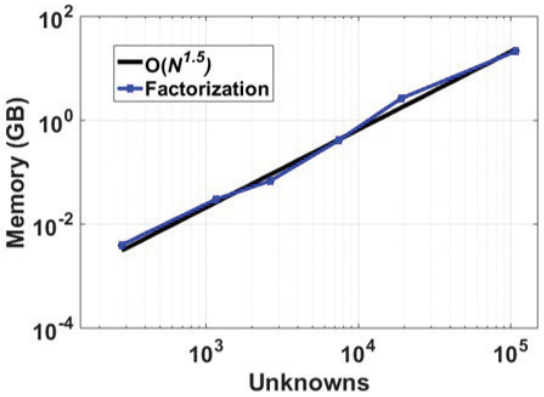


Fig. 6. Factorization memory with an increasing number of unknowns.

IV. NUMERICAL RESULTS

This section demonstrates the accuracy and efficiency of the proposed new fast direct solver method. The fast direct solver is applied on a single-level symmetric half-compressed block matrix computed with ACA compression tolerance of 1e-3 on 128 GB memory and an Intel (Xeon E5-2670) processor system. We have shown the accuracy of bistatic and monostatic RCS.

Efficiency is shown for the factorization and solution time for different geometries compared with open source fast direct solver code 3D EFIE ie3d code from ButterflyPACK [38] H-Matrix solver. All simulations were carried out using a double-precision data type.

A. Bistatic RCS

The accuracy of the proposed fast direct solver is shown for a 4λ sphere bistatic RCS. The bistatic RCS is computed for the plane wave incident at angle $\theta = 0^\circ$ and $\phi = 0^\circ$, with the observation angle varying from $\theta = 0^\circ$ to 180° and $\phi = 0^\circ$. The sphere meshes $\lambda/10$ edge length with 82,515 unknowns. The bistatic RCS is computed from the proposed fast direct solver ButterflyPACK and is compared with the Mie Series analytical method.

It is observed from Fig. 7 that the bistatic RCS computed from the fast direct solver matches well with the Mie Series analytical method.

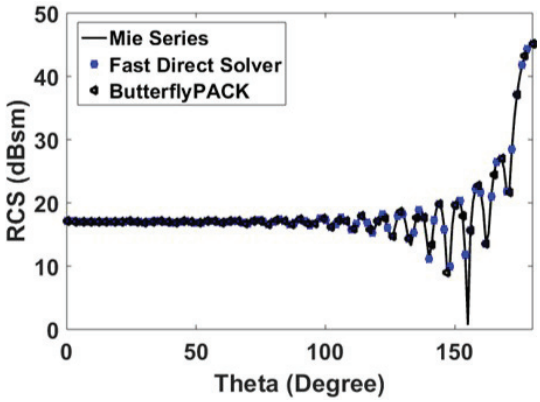


Fig. 7. Bistatic RCS for observation angle $\theta = 0^\circ$ to 180° and $\phi = 0^\circ$ with incident angle $\theta = 0^\circ$ and $\phi = 0^\circ$.

Table 1 shows the factorization and solution time and storage memory for the 4λ radius sphere from the proposed fast direct solver and open-source ButterflyPACK. It can be observed that the proposed fast direct solver is 1.3x faster than ButterflyPACK for 4λ sphere RCS computation.

Table 1: Factorization and solution time for 4λ sphere

	Fact. Time (s)	**Sol. Time (s)**	**Total Time (s)**	**Memory (GB)**
ButterflyPACK	26312.73	2.08	26314.81	21
Fast Direct Solver	20129.40	2.65	20132.05	27

B. Monostatic RCS

In this sub-section, we show the accuracy and efficiency of the proposed method for monostatic RCS computation, a multi-RHS problem. The computations are shown for an open and closed structure.

1. Square plate [36]

Monostatic RCS for 181 RHS is computed for the square plate of size $3\lambda \times 3\lambda$ with plane wave incident and observation angle varying from $\theta = -90^\circ$ to 90° from $\phi = 270^\circ$. The plate was meshed for $\lambda/10$ element size for 6033 unknowns. The RCS is computed using a ButterflyPACK solver and with the proposed fast direct solver. It is observed from Fig. 8 that the proposed method of RCS computation matches the open-source fast direct solver, and the results are given in [36].

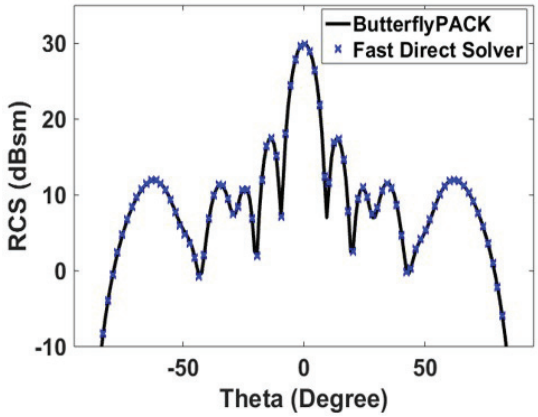


Fig. 8. Monostatic RCS of a square plate for a plane wave incident and observation angle varying from $\theta = -90^\circ$ to 90° and $\phi = 270^\circ$.

The factorization and solution time and storage memory in Table 2 are shown for the ButterflyPACK solver and the proposed fast direct solver for 181 RHS.

Table 2: Factorization and solution time for square plate

	Fact. Time (s)	**Sol. Time (s)**	**Total Time (s)**	**Memory (GB)**
ButterflyPACK	26.53	9.49	36.02	0.18
Fast Direct Solver	13.58	7.42	21.02	0.15

Table 2 shows the significant time saving for matrix factorization time and total time with the proposed fast direct solver. The proposed method is 1.7x faster than ButterflyPACK.

2. NASA almond [37]

Monostatic RCS for 180 RHS is computed for the NASA almond with the dimensions as in [35] at 7 GHz with a VV polarized plane wave incident and observation angle varying from $\phi = 0^\circ$ to 180° and $\theta^\circ = 90^\circ$. The CAD geometry is meshed for $\lambda/10$ element size for

8761 unknowns. The RCS is computed using the ButterflyPACK solver and with the proposed fast direct solver.

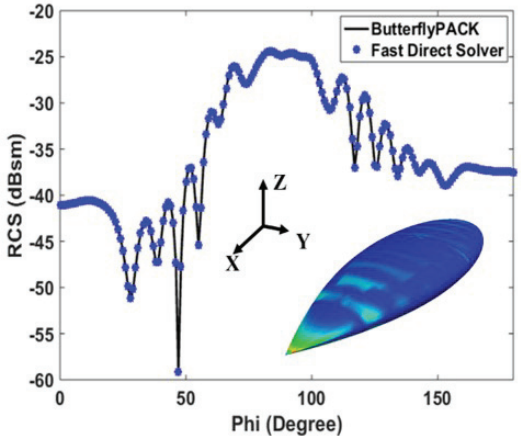


Fig. 9. Monostatic RCS of NASA almond for a plane wave incident and observation angle varying from $\phi = 0^\circ$ to 180° and $\theta^\circ = 90^\circ$.

It is observed that the proposed method of RCS computation matched with the ButterflyPACK solver, and the results are given in [37].

Table 3: Factorization and solution time for NASA almond

	Fact. Time (s)	**Sol. Time (s)**	**Total Time (s)**	**Memory (GB)**
ButterflyPACK	150.96	18.63	169.59	0.39
Fast Direct Solver	47.32	18.14	65.46	0.44

Matrix factorization and solution time and storage memory in Table 3 are shown for the ButterflyPACK solver and proposed fast direct solver for 180 RHS. The proposed fast direct solver is 2.5x faster than ButterflyPACK.

3. Ship

Monostatic RCS for 180 RHSs is computed for a ship of size 113 m length, 14 m width, and 20 m max height, and is computed at 80 MHz with a VV polarized plane wave. The incident and observation angles vary from $\phi = 0^\circ$ to 180° and $\theta^\circ = 90^\circ$. The geometry is meshed for $\lambda/10$ element size for 128,028 unknowns. RCS is computed using a multilevel ButterflyPACK solver and with the proposed fast direct solver.

It is observed from Fig. 10 that the proposed method of RCS computation matched with the ButterflyPACK solver.

The factorization time, solution time, and storage memory in Table 4 is shown for the ButterflyPACK solver and proposed fast direct solver for 180 RHS. It

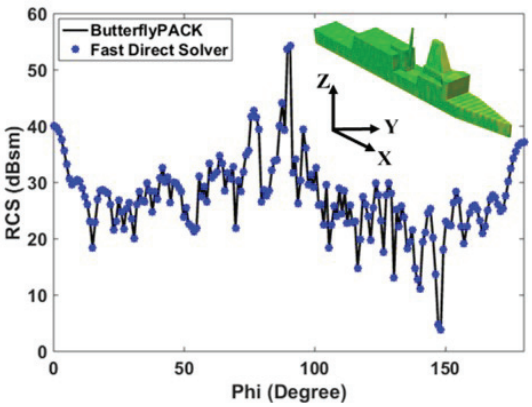


Fig. 10. Monostatic RCS of a ship-like object for a plane wave incident and observation angle varying from $\phi = 0^\circ$ to 180° and $\theta^\circ = 90^\circ$.

Table 4: Factorization and solution time for the ship

	Fact. Time (s)	**Sol. Time (s)**	**Total Time (s)**	**Memory (GB)**
ButterflyPACK	70694.09	772.30	71466.39	45
Fast Direct Solver	52391.67	1024.91	53416.58	51

is observed from Table 4 that there is a significant time saving for factorization time and total time with the proposed fast direct solver, and it is 1.3x faster than the ButterflyPACK solver.

V. CONCLUSION

The work proposed in this paper is a new fast direct solver based on the diagonalization method with the guaranteed solution for the MoM linear system of equations. The single-level symmetric half-compressed block matrix is factorized into $\boldsymbol{LDL'}$ where $\boldsymbol{D}$ is a diagonal block matrix, and $\boldsymbol{L'}$ is an upper triangle compressed block matrix. The high cost of diagonalization is overcome using a low-rank block matrix operation process. Our results show the factorization cost scaled to $O(N^{1.5})$. The solution depends on the block diagonal matrix solution process and matrix-vector product of the upper triangle $\boldsymbol{D}$. The far-field compressed matrices in the upper triangle blocks for each row block are merged further to save matrix-vector product cost and storage memory. Also, the solution process time for the proposed factorization scales to $O(N)$. In the current work, the matrix is compressed for the error tolerance of 1e-3. The factorization and solution cost can be reduced by compressing the matrix for lower error tolerance 1e-2. The low cost of the solution process, as demonstrated by illustrative examples, makes it highly desirable for solving multi-RHS problems like monostatic RCS computation or multiport network analysis. Unlike block $\boldsymbol{LU}$ factorization

and solution process, which is highly serial in nature, in the proposed method, the block operation can be done independently, making the process amendable for efficient parallelization for both factorization and solution.

REFERENCES

[1] R. F. Harrington, *Field Computation by Moment Methods*. Malabar, Florida: Krieger Publishing Co., 1982.

[2] J.-M. Jin, *The Finite Element Method in Electromagnetics*. Hoboken, New Jersey: John Wiley & Sons, 2015.

[3] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 1995.

[4] W. C. Chew, J.-M. Jin, E. Michielssen, and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*. Norwood, MA: Artech House, 2000.

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

[6] J. R. Phillips and J. K. White, "A pre-corrected FFT method for electrostatic analysis of complicated 3-D structures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 16, no. 10, pp. 1059-1072, Oct. 1997.

[7] M. Bebendorf, "Approximation of boundary element matrices," *Numerische Math*., vol. 86, no. 4, pp. 565-589, June 2000.

[8] S. Kurz, O. Rain, and S. Rjasanow, "The adaptive cross-approximation technique for the 3D boundary-element method," *IEEE Transaction on Magnetics,* vol. 38, no. 2, pp. 421-424, Mar. 2002.

[9] S. Kapur and D. E. Long, "IES3: Efficient electrostatic and electromagnetic solution," *IEEE Computer Science and Engineering*, vol. 5, no. 4, pp. 60-66, Dec. 1998.

[10] W. Hackbusch, "A sparse matrix arithmetic based on H-matrices. Part I: Introduction to H-matrices," *Computing*, vol. 62, no. 2, pp. 89-108, Apr. 1999.

[11] W. Hackbusch and B. N. Khoromskij, "A sparse H-Matrix arithmetic. Part II: Application to multi-dimensional problems," *Computing*, vol. 64, pp. 21-47, Feb. 2000.

[12] Y. K. Negi, "Memory reduced half hierarchal matrix (H-Matrix) for electrodynamic electric field integral equation," *Progress in Electromagnetics Research Letters*, vol. 96, pp. 91-96, Feb. 2021.

[13] A. F. Peterson, "The 'interior resonance' problem associated with surface integral equations of electromagnetics: Numerical consequences and a survey of remedies," *Electromagnetics*, vol. 10, no. 3, pp. 293-312, July 1990.

[14] R. E. Hodges and Y. Rahmat-Samii, "The evaluation of MFIE integrals with the use of vector triangle basis functions," *Microwave and Optical Technology Letters*, vol. 14, no. 1, pp. 9-14, Jan. 1997.

[15] P. Yla-Oijala and M. Taskinen, "Application of combined field integral equation for electromagnetic scattering by dielectric and composite objects," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 3, pp. 1168-1173, Mar. 2005.

[16] Y. Saad, "ILUT: A dual threshold incomplete *LU* factorization," *Numerical Linear Algebra Application*, vol. 1, no. 4, pp. 387-402, 1994.

[17] Y. K. Negi, N. Balakrishnan, Sadasiva M. Rao, and Dipanjan Gope, "Null-field preconditioner with selected far-field contribution for 3-D full-wave EFIE," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 11, pp. 4923-4928, Aug. 2016.

[18] Y. K. Negi, N. Balakrishnan, and Sadasiva M. Rao, "Symmetric near-field Schur's complement preconditioner for hierarchal electric field integral equation solver," *IET Microwaves, Antennas & Propagation*, vol. 14, no. 14, pp. 1846-1856, Nov. 2020.

[19] D. Gope, I. Chowdhury, and V. Jandhyala, "DiMES: Multilevel fast direct solver based on multipole expansions for parasitic extraction of massively coupled 3D microelectronic structures," in *Proceedings of the 42nd annual Design Automation Conference*, Anaheim, CA, USA, pp. 159-162, June 2005.

[20] D. Sushnikova, L. Greengard, M. O'Neil, and M. Rachh, "FMM-LU: A fast direct solver for multiscale boundary integral equations in three dimensions," *Multiscale Modeling & Simulation*, vol. 21, no. 4, pp. 1570-1601, Dec. 2023.

[21] L. Greengard, D. Gueyffier, P.-G. Martinsson, and V. Rokhlin, "Fast direct solvers for integral equations in complex three-dimensional domains," *Acta Numerica*, vol. 18, pp. 243-275, May 2009.

[22] M. Ma and D. Jiao, "Accuracy directly controlled fast direct solution of general H∧2-Matrices and its application to solving electrodynamic volume integral equations," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 1, pp. 35-48, Jan. 2018.

[23] Y. K. Negi, N. Balakrishnan, and S. M. Rao, "Fast power series solution of large 3-D electrodynamic integral equation for PEC scatterers," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 36, no. 10, pp 1301-1311, Oct. 2021.

[24] Y. K. Negi, N. Balakrishnan, and S. M. Rao, "Multilevel power series solution for large surface and volume electric field integral equation," *Applied Computational Electromagnetics Society (ACES) Journal,* vol. 38, no. 5, pp. 297-303, Sep. 2023.

[25] S. Huang and Y. J. Liu, "A new fast direct solver for the boundary element method," *Computational Mechanics*, vol. 60, pp. 379-392, Sep. 2017.

[26] Z. Rong, M. Jiang, Y. Chen, L. Lei, X. Li, and Z. Nie, "Fast direct solution of integral equations with modified HODLR structure for analyzing electromagnetic scattering problems," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 5, pp. 3288-3296, May 2019.

[27] A. Heldring, J. M. Rius, J. M. Tamayo, J. Parron, and E. Ubeda, "Fast direct solution of Method of Moments linear system," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 11, pp. 3220-3228, Nov. 2007.

[28] J. Shaeffer, "Direct solve of electrically large integral equations for problem sizes to 1 M unknowns," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 8, pp. 2306-2313, Aug. 2008.

[29] W. C. Gibson, "Efficient solution of electromagnetic scattering problems using multilevel adaptive cross approximation and *LU* factorization," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 5, pp. 3815-3823, May 2020.

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

[31] Y. K. Negi, V. P. Padhy, and N. Balakrishnan, "Re-compressed H-Matrices for fast electric field integral equation," in *IEEE-International Conference on Computational Electromagnetics (ICCEM 2020),* Singapore, pp. 24-26, Aug. 2020.

[32] M. Neumann, "On the Schur complement and the LU-factorization of a matrix," *Linear and Multilinear Algebra*, vol. 9, no. 4, pp. 241-254, Jan. 1981.

[33] S. M. Rao, "A true domain decomposition procedure based on Method of Moments to handle electrically large bodies," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 9, pp. 4233-4238, Sep. 2012.

[34] Y. K. Negi, N. Balakrishnan, and S. M. Rao, "Schur decomposition fast direct solver for volume surface integral equation," in *2024 IEEE International Symposium on Antennas and Propagation and INC/USNC-URSI Radio Science Meeting (AP-S/INC-USNC-URSI)*, Firenze, Italy, pp. 973-974, July 2024.

[35] J. Shaeffer, "Low-rank matrix algebra for the Method of Moments," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 33, no. 10, pp. 1052-1059, Oct. 2018.

[36] K. L. Virga and Y. Rahmat-Samii, "RCS characterization of a finite ground plane with perforated apertures: Simulations and measurements," *IEEE Transactions on Antennas and Propagation*, vol. 42, no. 11, pp. 1491-1501, Nov. 1994.

[37] A. C. Woo, H. T. G. Wang, M. J. Schuh, and M. L. Sanders, "EM programmer's notebook-benchmark radar targets for the validation of computational electromagnetics programs," *IEEE Antennas and Propagation Magazine*, vol. 35, no. 1, pp. 84-89, Feb. 1993.

[38] L. Y. Zhuan, *ButterflyPACK* [Online]. Available: https://github.com/liuyangzhuan/ButterflyPACK2

Yoginder Kumar Negi obtained the B.Tech. degree in Electronics and Communication Engineering from Guru Gobind Singh Indraprastha University, New Delhi, India, in 2005, M.Tech. degree in Microwave Electronics from Delhi University, New Delhi, India, in 2007, and the Ph.D. degree in engineering from Indian Institute of Science (IISc), Bangalore, India, in 2018. Negi joined Supercomputer Education Research Center (SERC), IISc Bangalore, in 2008 as a Scientific Officer. He is currently working as a Senior Scientific Officer in SERC IISc Bangalore. His current research interests include numerical electromagnetics, fast techniques for electromagnetic application, bio-electromagnetics, high-performance computing, and antenna design and analysis.

N. Balakrishnan received the B.E. degree (Hons.) in Electronics and Communication from the University of Madras, Chennai, India, in 1972, and the Ph.D. degree from the Indian Institute of Science, Bengaluru, India, in 1979. He joined the Department of Aerospace Engineering, Indian Institute of Science, as an Assistant Professor, in 1981, where he became a Full Professor in 1991, served as Associate Director from 2005 to 2014, and is currently an INSA Senior Scientist at the Supercomputer Education and Research Centre. He has authored over 200 publications in international journals and international conferences. His current research interests include numerical electromagnetics,

high-performance computing and networks, polarimetric radars and aerospace electronic systems, information security, and digital library. N. Balakrishnan is a fellow of The World Academy of Sciences (TWAS), the National Academy of Science, the Indian Academy of Sciences, the Indian National Academy of Engineering, the National Academy of Sciences, and the Institution of Electronics and Telecommunication Engineers.

Sadasiva M. Rao obtained his Bachelors, Masters, and Doctoral degrees in electrical engineering from Osmania University, Hyderabad, India, Indian Institute of Science, Bangalore, India, and University of Mississippi, USA, in 1974, 1976, and 1980, respectively. He is well known in the electromagnetic engineering community and included in Thomson Scientifics' *Highly Cited Researchers List*. Rao has been teaching electromagnetic theory, communication systems, electrical circuits, and other related courses at the undergraduate and graduate level for the past 30 years at various institutions. At present, he is working at the Naval Research Laboratories, USA. He has published/presented over 200 papers in various journals/conferences. He is an elected Fellow of IEEE.

Surface Integral Equations in Computational Electromagnetics: A Comprehensive Overview of Theory, Formulations, Discretization Schemes and Implementations

Parmenion S. Mavrikakis and Olivier J. F. Martin

Nanophotonics and Metrology Laboratory
École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
parmenion.mavrikakis@epfl.ch, olivier.martin@epfl.ch

Abstract **–** Computational electromagnetics based on surface integral equations provides accurate and efficient solutions for three-dimensional electromagnetic scattering problems in the frequency domain. In this review paper, we first introduce a complete and detailed theoretical analysis of the surface integral equation method, including different properties of the corresponding integral operators and equations. Using a pedagogical approach that should appeal to electrical engineers, we provide a systematic and comprehensive derivation of the different formulations found in the literature and discuss their advantages and pitfalls. Additionally, we provide a mathematical overview of the corresponding function spaces that clarifies the importance of correctly combining basis and testing functions and we examine the various aspects of discretization schemes, such as the Green's function singularity subtraction and the application of different testing methods. Moreover, we assess alternative formulations and discretization procedures and draw particular conclusions about them, by comparing numerous examples and results from previously published works. Finally, we provide a detailed discussion on numerical solvers and approaches.

Index Terms **–** Basis function, Buffa-Christiansen function, Calderon preconditioner, discretization, electromagnetic scattering, fast solvers, formulation, Green's function, half-Rao-Wilton-Glisson function, high-performance computing, integral equation, integral operator, low frequency breakdown, Method of Moments, Rao-Wilton-Glisson function, singularity subtraction, surface integral equation, testing function, testing method, Trintinalia-Ling functions.

I. INTRODUCTION

For many decades, computational electromagnetics has been playing a crucial role both in academia and in industry, to investigate a plethora of phenomena [1–6]. Indeed, analytical expressions for electromagnetic scattering problems exist only for objects with very specific shapes. For scattering, Mie theory can be used only to model analytically homogeneous, coated, or multilayered spheres and infinite cylinders [7]. The treatment of scatterers with more complex geometries requires the use of accurate numerical techniques.

A very popular technique used for modeling scattering objects is the discrete dipole approximation (DDA), which approximates the scatterer by a finite array of polarizable point dipoles [8, 9], but presents enormous computational costs for large objects [10]. Numerical techniques that are based on differential formulations like the finite element method (FEM) [11–13] and the finite-difference time-domain (FDTD) [14–16] have sparse matrices in the final linear system, which is a significant advantage since they can use efficient storage and solver algorithms with better performances. Nevertheless, the discretization of the whole computational domain (including the surrounding medium) and the existence of spurious modes represent some major concerns for these approaches [17].

Volume integral equations (VIEs) use the integral form of Maxwell's equations [18]. Contrary to differential techniques, only the discretization of the scatterer's domain is needed [19]. The surrounding is not discretized since its effect is intrinsically included in the formulation. Also, VIEs can be used to model inhomogeneous materials. However, an important disadvantage of VIEs is that they involve densely populated matrices [20].

Surface integral equations (SIEs), are very efficient because they only require the discretization of the scatterer's surface, thus presenting a significantly smaller number of unknowns compared to VIEs [21]. However, they are limited to piecewise homogeneous materials since they use the Green's function of a homogeneous region. The SIE method is a widely used numerical approach for analyzing electromagnetic scattering in metallic, dielectric, and composite metallic-dielectric structures [22]. SIEs can be formulated in many different ways for the same electromagnetic problem. Thus, one may wonder what is the best formulation to use?

Submitted On: October 9, 2024
Accepted On: April 17, 2025

https://doi.org/10.13052/2025.ACES.J.400402

Some of the most popular SIE formulations for impenetrable scatterers are the electric field integral equation (EFIE), magnetic field integral equation (MFIE) and combined-field integral equation (CFIE) [23]. For penetrable objects the most used formulations are the Poggio-Miller-Chang-Harrington-Wu-Tsai (PMCHWT) [24] and Müller [25]. In the 1970s and 1980s many numerical approaches were introduced for solving the aforementioned equations, initially for rotationally symmetric objects [23, 26, 27], and subsequently for arbitrarily shaped three-dimensional (3D) objects [28–30]. SIEs can be categorized into the Fredholm integral equations of the first and the second kinds [31]. EFIE and PMCHWT equations are both of the first kind. MFIE and Müller equations belong to the second kind. The first kind of surface integral equations present a superior numerical accuracy compared to the second kind, while the latter show better performance regarding iterative solutions and convergence [32]. Hence, identifying a specific formulation as the most optimal is not straightforward. The main reason behind this trade-off between accuracy and iterative solution convergence has been proven to originate from the improper testing of the second kind of formulations [33, 34]. Hence, in every problem, the discretization scheme (selection of basis and testing functions) and the SIE formulation must be selected carefully, since the numerical accuracy and efficiency of the solution are heavily affected by this formulation-discretization combination.

In this review, we present the theoretical analysis of a general electromagnetic problem that leads to different SIE formulations. We discuss the integral operators, the integral equations and the different formulations. Moreover, we delve into the various components of a discretization scheme, such as the singularity subtraction and the utilization of different basis and testing functions, by providing also a mathematical overview of function spaces to explain the significance of combining basis and testing functions properly. Furthermore, we go through several results from various published works, in order to compare different formulations and discretization procedures, so that we can draw specific conclusions about them alone and also about their combination. Finally, we provide a detailed discussion on numerical solvers and approaches.

II. THEORETICAL FRAMEWORK

A. Scattering problems

Consider Fig. 1, where a scatterer is placed in a homogeneous background. The scatterer is of arbitrary shape and consists of piecewise homogeneous media. Every region is a domain Ω_i with a constant electric permittivity ε_i and magnetic permeability μ_i. The time dependence $e^{-j\omega t}$ is used throughout.

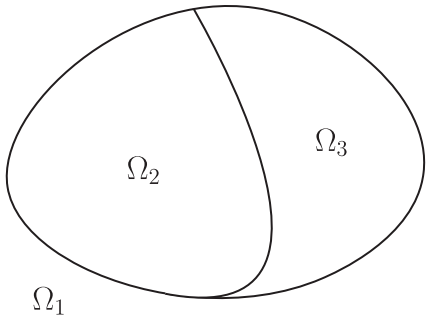


Fig. 1. Scatterer in a homogeneous background Ω_1.

For each domain,

$$\nabla\times\nabla\times\mathbf{E}_i(\mathbf{r})-k_i^2\mathbf{E}_i(\mathbf{r})=j\omega\mu_i\mathbf{J}_i(\mathbf{r}), \tag{1}$$

where $k_i=\omega\sqrt{\varepsilon_i\mu_i}$. The Dyadic Green's function $\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')$ solves the equation [35]

$$\nabla\times\nabla\times\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')-k_i^2\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')=\overline{\mathbf{1}}\delta(\mathbf{r}-\mathbf{r}'). \tag{2}$$

The multiplication of (1) with $\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')$ from the right and (2) with $\overline{\mathbf{E}}_i(\mathbf{r})$ from the left gives

$$\begin{aligned}&\nabla\times\nabla\times\mathbf{E}_i(\mathbf{r})\cdot\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')-\\&\mathbf{E}_i(\mathbf{r})\cdot\nabla\times\nabla\times\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')=\\&j\omega\mu_i\mathbf{J}_i(\mathbf{r})\cdot\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')-\mathbf{E}_i(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}').\end{aligned} \tag{3}$$

Integrating (3) over Ω_i and using the following relation [36]

$$\begin{aligned}&\nabla\cdot\Big([\nabla\times\mathbf{E}(\mathbf{r})]\times\overline{\mathbf{G}}(\mathbf{r},\mathbf{r}')+\\&\mathbf{E}(\mathbf{r})\times[\nabla\times\overline{\mathbf{G}}(\mathbf{r},\mathbf{r}')]\Big)=\\&\nabla\times\nabla\times\mathbf{E}(\mathbf{r})\cdot\overline{\mathbf{G}}(\mathbf{r},\mathbf{r}')-\mathbf{E}(\mathbf{r})\cdot\nabla\times\nabla\times\overline{\mathbf{G}}(\mathbf{r},\mathbf{r}'),\end{aligned} \tag{4}$$

gives

$$\begin{aligned}&\int_{\Omega_i}\mathrm{d}V\nabla\cdot\Big([\nabla\times\mathbf{E}_i(\mathbf{r})]\times\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')+\\&\mathbf{E}_i(\mathbf{r})\times[\nabla\times\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')]\Big)=\\&\mathbf{E}_i^{\mathrm{inc}}(\mathbf{r}')-\begin{cases}\mathbf{E}_i(\mathbf{r}'), & \mathbf{r}'\in\Omega_i\\ \kappa(\mathbf{r}')\mathbf{E}_i(\mathbf{r}'), & \mathbf{r}'\in\partial\Omega_i\\ 0, & \mathbf{r}'\notin\Omega_i\end{cases},\end{aligned} \tag{5}$$

where

$$\begin{aligned}\mathbf{E}_i^{\mathrm{inc}}(\mathbf{r}')&=j\omega\mu_i\int_{\Omega_i}\mathrm{d}V\mathbf{J}_i(\mathbf{r})\cdot\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\\&=j\omega\mu_i\int_{\Omega_i}\mathrm{d}V\overline{\mathbf{G}}_i(\mathbf{r}',\mathbf{r})\cdot\mathbf{J}_i(\mathbf{r}),\end{aligned} \tag{6}$$

is the incident electric field intensity generated by the current density $\mathbf{J}_i(\mathbf{r})$ inside Ω_i. Regarding the transposition of the Dyadic Green's function [37], $\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')^T=\overline{\mathbf{G}}_i(\mathbf{r}',\mathbf{r})$. Next, by using Gauss' theorem, the following

surface integral emerges

$$\int_{\partial\Omega_i} \mathrm{d}S\hat{\mathbf{n}}_i \cdot \Big([\nabla \times \mathbf{E}_i(\mathbf{r})] \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') + \mathbf{E}_i(\mathbf{r}) \times \left[\nabla \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] \Big) = \mathbf{E}_i^{\text{inc}}(\mathbf{r}') - \begin{cases} \mathbf{E}_i(\mathbf{r}'), & \mathbf{r}' \in \Omega_i \\ \kappa(\mathbf{r}')\mathbf{E}_i(\mathbf{r}'), & \mathbf{r}' \in \partial\Omega_i \\ 0, & \mathbf{r}' \notin \Omega_i \end{cases}, \quad (7)$$

where $\hat{\mathbf{n}}_i$ is the unit normal vector on $\partial\Omega_i$ with direction from the inside to the outside of Ω_i. By using the time-harmonic nature of the different field quantities, the kernel of the integral becomes [37]

$$\begin{aligned} &\hat{\mathbf{n}}_i(\mathbf{r}) \cdot [\nabla \times \mathbf{E}_i(\mathbf{r})] \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') = \\ &\hat{\mathbf{n}}_i(\mathbf{r}) \times [\nabla \times \mathbf{E}_i(\mathbf{r})] \cdot \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') = \\ &j\omega\mu_i \overline{\mathbf{G}}_i(\mathbf{r}',\mathbf{r}) \cdot [\hat{\mathbf{n}}_i(\mathbf{r}) \times \mathbf{H}_i(\mathbf{r})], \end{aligned} \quad (8)$$

and with the use of

$$\left[\nabla \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right]^{\mathbf{T}} = -\nabla \times \overline{\mathbf{G}}_i(\mathbf{r}',\mathbf{r}), \quad (9)$$

the second kernel term becomes

$$\begin{aligned} &\hat{\mathbf{n}}_i(\mathbf{r}) \cdot \mathbf{E}_i(\mathbf{r}) \times \left[\nabla \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] = \\ &[\hat{\mathbf{n}}_i(\mathbf{r}) \times \mathbf{E}_i(\mathbf{r})] \cdot \left[\nabla \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] = \\ &-\left[\nabla \times \overline{\mathbf{G}}_i(\mathbf{r}',\mathbf{r})\right] \cdot [\hat{\mathbf{n}}_i(\mathbf{r}) \times \mathbf{E}_i(\mathbf{r})]. \end{aligned} \quad (10)$$

Finally, by introducing the surface electric and magnetic current densities

$$\mathbf{J}_{s,i}(\mathbf{r}) = \hat{\mathbf{n}}_i(\mathbf{r}) \times \mathbf{H}_i(\mathbf{r}) \quad (11)$$

$$\mathbf{M}_{s,i}(\mathbf{r}) = -\hat{\mathbf{n}}_i(\mathbf{r}) \times \mathbf{E}_i(\mathbf{r}), \quad (12)$$

where $\hat{\mathbf{n}}_i(\mathbf{r})$ is the unit normal vector of $\partial\Omega_i$ towards the outer side of Ω_i. Equation (7) becomes [18]

$$\eta_i \left[jk_i \int_{\partial\Omega_i} \mathrm{d}S' \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_{s,i}(\mathbf{r}') \right] + \int_{\partial\Omega_i} \mathrm{d}S' \left[\nabla' \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] \cdot \mathbf{M}_{s,i}(\mathbf{r}') = \mathbf{E}_i^{\text{inc}}(\mathbf{r}) - \begin{cases} \mathbf{E}_i(\mathbf{r}), & \mathbf{r} \in \Omega_i \\ \kappa(\mathbf{r})\mathbf{E}_i(\mathbf{r}), & \mathbf{r} \in \partial\Omega_i \\ 0, & \mathbf{r} \notin \Omega_i \end{cases}, \quad (13)$$

where $\mathbf{r}$ and $\mathbf{r}'$ have been swapped. Also, $\kappa(\mathbf{r}) = 1 - \Omega(\mathbf{r})/4\pi$, where $\Omega(\mathbf{r})$ is the solid angle subtended by the observation point $\mathbf{r}$ [18]. For locally smooth surfaces $\Omega(\mathbf{r}) = 2\pi$, thus $\kappa(\mathbf{r}) = 1/2$.

A similar analysis can be applied to the case of the magnetic field [38]. Starting from the following equation

$$\nabla \times \nabla \times \mathbf{H}_i(\mathbf{r}) - k_i^2 \mathbf{H}_i(\mathbf{r}) = \nabla \times \mathbf{J}_i(\mathbf{r}). \quad (14)$$

By identifying the incident magnetic field intensity as

$$\begin{aligned} \mathbf{H}_i^{\text{inc}}(\mathbf{r}') &= \int_{\Omega_i} \mathrm{d}V \mathbf{J}_i(\mathbf{r}) \cdot \left[\nabla' \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] \\ &= \int_{\Omega_i} \mathrm{d}V \overline{\mathbf{G}}_i(\mathbf{r}',\mathbf{r}) \cdot [\nabla \times \mathbf{J}_i(\mathbf{r})], \end{aligned} \quad (15)$$

an analogous equation is derived

$$\eta_i^{-1} \left[jk_i \int_{\partial\Omega_i} \mathrm{d}S' \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') \cdot \mathbf{M}_{s,i}(\mathbf{r}') \right] - \int_{\partial\Omega_i} \mathrm{d}S' \left[\nabla' \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] \cdot \mathbf{J}_{s,i}(\mathbf{r}') = \mathbf{H}_i^{\text{inc}}(\mathbf{r}) - \begin{cases} \mathbf{H}_i(\mathbf{r}), & \mathbf{r} \in \Omega_i \\ \kappa(\mathbf{r})\mathbf{H}_i(\mathbf{r}), & \mathbf{r} \in \partial\Omega_i \\ 0, & \mathbf{r} \notin \Omega_i \end{cases}. \quad (16)$$

After the solution of the above integral equations, the currents can be used to calculate the fields at any position $\mathbf{r} \in \Omega_i$, as follows

$$\mathbf{E}_i(\mathbf{r}) = \mathbf{E}_i^{\text{inc}}(\mathbf{r}) - \eta_i \left[jk_i \int_{\partial\Omega_i} \mathrm{d}S' \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_{s,i}(\mathbf{r}') \right] - \int_{\partial\Omega_i} \mathrm{d}S' \left[\nabla' \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] \cdot \mathbf{M}_{s,i}(\mathbf{r}'), \quad (17)$$

$$\mathbf{H}_i(\mathbf{r}) = \mathbf{H}_i^{\text{inc}}(\mathbf{r}) - \eta_i^{-1} \left[jk_i \int_{\partial\Omega_i} \mathrm{d}S' \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') \cdot \mathbf{M}_{s,i}(\mathbf{r}') \right] + \int_{\partial\Omega_i} \mathrm{d}S' \left[\nabla' \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] \cdot \mathbf{J}_{s,i}(\mathbf{r}'). \quad (18)$$

B. Boundary conditions

As illustrated in Fig. 2, we assume the existence of two domains Ω_1 and Ω_2 with different media, and a boundary $\partial\Omega = \partial\Omega_1 \cap \partial\Omega_2$. Maxwell's equations require that the tangential components of the electric and magnetic fields are continuous across the boundary $\partial\Omega$, as depicted in the following equations, where $\mathbf{r} \in \partial\Omega$:

$$\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}_1(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}_2(\mathbf{r}), \quad (19)$$

$$-\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}_1(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}_2(\mathbf{r}), \quad (20)$$

where $\hat{\mathbf{n}}(\mathbf{r}) = \hat{\mathbf{n}}_1(\mathbf{r})$ is the the unit normal vector of $\partial\Omega$ pointing towards Ω_2 and $\hat{\mathbf{n}}_2(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r})$, as shown in Fig. 2. Thus, when it comes to the surface electric and magnetic current densities on $\partial\Omega_1$ and $\partial\Omega_2$, the relation between them for $\mathbf{r} \in \partial\Omega$ is

$$\mathbf{J}_{s,1}(\mathbf{r}) = -\mathbf{J}_{s,2}(\mathbf{r}), \quad (21)$$

$$\mathbf{M}_{s,1}(\mathbf{r}) = -\mathbf{M}_{s,2}(\mathbf{r}). \quad (22)$$

The above relations between surface current densities play a key role not only for the theoretical formulation, but also for its discretization.

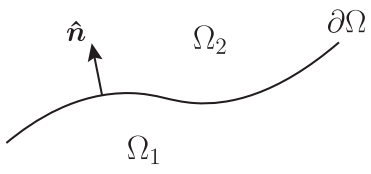


Fig. 2. Boundary between the domains Ω_1, Ω_2, and definition of the normal $\hat{\mathbf{n}}(\mathbf{r})$.

III. SURFACE INTEGRAL FORMULATIONS

A. Integral operators

Formally, a linear integral equation can be written

$$\mathscr{L}\{\mathbf{s}\} = \mathbf{R}, \tag{23}$$

where $\mathscr{L}$ is the integral (linear) operator, $\mathbf{s}$ is the unknown quantity (the electric and magnetic surface current densities) and $\mathbf{R}$ is the known quantity (the excitation). The surface integral operator

$$\mathscr{L}\{\mathbf{s}\}(\mathbf{r}) = \int_{\partial\Omega} \mathrm{d}S' K(\mathbf{r},\mathbf{r}')\,\mathbf{s}(\mathbf{r}') \tag{24}$$

includes a kernel $K(\mathbf{r},\mathbf{r}')$ and the integral runs over a boundary $\partial\Omega$ of the geometry.

The linear integral operator can be categorized according to its singularity. If the singularity order is less than the integral's dimension, then the operator is weakly singular [39]. An example of such an operator is the following

$$\mathscr{L}_w\{\mathbf{s}\}(\mathbf{r}) = \int_{\partial\Omega} \mathrm{d}S' \frac{1}{|\mathbf{r}-\mathbf{r}'|}\mathbf{s}(\mathbf{r}') . \tag{25}$$

In this case, as $|\mathbf{r}-\mathbf{r}'| \to 0$, the kernel becomes singular. Its singularity dimension is equal to 1. However, since we are integrating over the 2D surface $\partial\Omega$, the integral's dimension is equal to 2. Thus, this is a weak (or mild) singularity, and the integral remains finite. Additionally, a weakly singular integral operator $\mathscr{L}_w$ is bounded and maps a function to a smoother one, because its range space is reduced by one order relative to its original domain [40]. Furthermore, the spectrum of a bounded operator accumulates to a constant and, in the special case that it accumulates to zero, the integral operator is said to be compact [41]. Regarding the previously presented example, the integral operator $\mathscr{L}_w$ is compact, meaning its eigenvalues (in spectral terms) accumulate to zero, and the operator tends to smooth the function it acts upon. If the singularity order is equal to or larger than the integral's dimension, then the operator is singular or hyper-singular, respectively. A typical hyper-singular example arises from increasing the power of the kernel's denominator, as follows

$$\mathscr{L}_h\{\mathbf{s}\}(\mathbf{r}) = \int_{\partial\Omega} \mathrm{d}S' \frac{1}{|\mathbf{r}-\mathbf{r}'|^3}\mathbf{s}(\mathbf{r}') . \tag{26}$$

In this scenario, as $|\mathbf{r}-\mathbf{r}'| \to 0$, the kernel becomes singular. However, its singularity dimension is equal to 3 and it is larger than the integral's dimension. Operators with (hyper-)singular kernels can lead to the appearance of an unbounded operator, with a spectrum that tends to go to infinity [42]. In the following sections we identify the integral operators that appear in (13) and (16) of the aforementioned analysis.

1. $\mathscr{D}$ operator

In (13) and (16) there is the linear integral operator

$$\mathscr{D}_i\{\mathscr{X}\}(\mathbf{r}) = jk_i \int_{\partial\Omega_i} \mathrm{d}S' \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') \cdot \mathscr{X}(\mathbf{r}') , \tag{27}$$

where $\mathscr{X}(\mathbf{r})$ can be the surface electric current density $\mathbf{J}_{s,i}(\mathbf{r})$ or the surface magnetic current density $\mathbf{M}_{s,i}(\mathbf{r})$. By taking into consideration that

$$\begin{aligned}\overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') &= \left(\overline{\mathbf{1}} + \frac{\nabla\nabla}{k_i^2}\right)\frac{e^{jk_i|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \\ &= \left(\overline{\mathbf{1}} + \frac{\nabla\nabla}{k_i^2}\right) G_i(\mathbf{r},\mathbf{r}') ,\end{aligned} \tag{28}$$

equation (27) can lead to the following calculations

$$\begin{aligned}\mathscr{D}_i\{\mathscr{X}\}(\mathbf{r}) &= jk_i \int_{\partial\Omega_i} \mathrm{d}S' \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') \cdot \mathscr{X}(\mathbf{r}') \\ &= jk_i \int_{\partial\Omega_i} \mathrm{d}S' G_i(\mathbf{r},\mathbf{r}')\,\mathscr{X}(\mathbf{r}') - \\ &\quad \frac{1}{jk_i}\nabla \int_{\partial\Omega_i} \mathrm{d}S' G_i(\mathbf{r},\mathbf{r}')\,\nabla_s' \cdot \mathscr{X}(\mathbf{r}') .\end{aligned} \tag{29}$$

2. $\mathscr{K}$ operator

In (13) and (16) we can also identify the linear integral operator

$$\mathscr{K}_i\{\mathscr{X}\}(\mathbf{r}) = \int_{\partial\Omega_i} \mathrm{d}S' \left[\nabla' \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] \cdot \mathscr{X}(\mathbf{r}') , \tag{30}$$

which, with the use of equation $\nabla' \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}') = \nabla' G_i(\mathbf{r},\mathbf{r}') \times \overline{\mathbf{1}}$, can be further written as follows

$$\begin{aligned}\mathscr{K}_i\{\mathscr{X}\}(\mathbf{r}) &= \int_{\partial\Omega_i} \mathrm{d}S' \left[\nabla' \times \overline{\mathbf{G}}_i(\mathbf{r},\mathbf{r}')\right] \cdot \mathscr{X}(\mathbf{r}') \\ &= \int_{\partial\Omega_i} \mathrm{d}S' \left[\nabla' G_i(\mathbf{r},\mathbf{r}')\right] \times \mathscr{X}(\mathbf{r}') .\end{aligned} \tag{31}$$

B. Integral equations

As mentioned in section I, two different kinds of integral equations exist. The integral equations of the first kind,

$$\mathscr{L}_1\{\mathbf{s}\} = \mathbf{R}, \tag{32}$$

and those of the second kind,

$$(\mathscr{L}_2 + \mathscr{I})\{\mathbf{s}\} = \mathbf{R}, \tag{33}$$

where $\mathscr{I}\{\mathbf{s}\}$ is the identity operator [41]. An integral equation of the first kind has a unique solution if the linear integral operator is coercive and one-to-one [43]. For an integral equation of the second kind, a unique solution exists when the operator $\mathscr{L}_2 + \mathscr{I}$ is one-to-one and the integral operator $\mathscr{L}_2$ is compact [43]. In general, an integral equation with an operator of the form $\mathscr{L} + \mathscr{K}$ has a unique solution if $\mathscr{L} + \mathscr{K}$ is one-to-one, $\mathscr{L}$ is compact, and $\mathscr{K}$ is a bounded operator with a bounded inverse [42–44]. Hence, in order to have unique solutions, in (32) the linear integral operator should not be compact, and in (33) it should not be unbounded [42].

According to the calculations of the previous sections, (13) and (16) can be rewritten on a (locally)

smooth surface $\partial\Omega_i$ as follows

$$\eta_i \mathscr{D}_i \{\mathbf{J}_{s,i}\}(\mathbf{r}) + \mathscr{K}_i \{\mathbf{M}_{s,i}\}(\mathbf{r}) = \mathbf{E}_i^{\text{inc}}(\mathbf{r}) - \frac{1}{2}\mathbf{E}_i(\mathbf{r}), \quad \mathbf{r} \in \partial\Omega_i, \tag{34}$$

$$-\mathscr{K}_i \{\mathbf{J}_{s,i}\}(\mathbf{r}) + \eta_i^{-1} \mathscr{D}_i \{\mathbf{M}_{s,i}\}(\mathbf{r}) = \mathbf{H}_i^{\text{inc}}(\mathbf{r}) - \frac{1}{2}\mathbf{H}_i(\mathbf{r}), \quad \mathbf{r} \in \partial\Omega_i. \tag{35}$$

The electric and magnetic fields can then be split into components parallel and perpendicular to the boundary $\partial\Omega_i$:

$$\mathbf{E}_i(\mathbf{r}) = [\mathbf{E}_i(\mathbf{r}) \times \hat{\mathbf{n}}_i(\mathbf{r})] \times \hat{\mathbf{n}}_i(\mathbf{r}) + [\hat{\mathbf{n}}_i(\mathbf{r}) \cdot \mathbf{E}_i(\mathbf{r})]\hat{\mathbf{n}}_i(\mathbf{r}), \tag{36}$$

$$\mathbf{H}_i(\mathbf{r}) = [\mathbf{H}_i(\mathbf{r}) \times \hat{\mathbf{n}}_i(\mathbf{r})] \times \hat{\mathbf{n}}_i(\mathbf{r}) + [\hat{\mathbf{n}}_i(\mathbf{r}) \cdot \mathbf{H}_i(\mathbf{r})]\hat{\mathbf{n}}_i(\mathbf{r}). \tag{37}$$

Using the boundary conditions and the continuity equations leads to the following expressions for (34) and (35):

$$\eta_i \mathscr{D}\{\mathbf{J}_{s,i}\}(\mathbf{r}) + \mathscr{K}\{\mathbf{M}_{s,i}\}(\mathbf{r}) - \frac{1}{2}\hat{\mathbf{n}}_i(\mathbf{r}) \times \mathbf{M}_{s,i}(\mathbf{r}) + j\frac{\eta_i}{2k_i}[\nabla_s \cdot \mathbf{J}_{s,i}(\mathbf{r})]\hat{\mathbf{n}}_i(\mathbf{r}) = \mathbf{E}_i^{\text{inc}}(\mathbf{r}), \mathbf{r} \in \partial\Omega_i, \tag{38}$$

$$-\mathscr{K}\{\mathbf{J}_{s,i}\}(\mathbf{r}) + \eta_i^{-1}\mathscr{D}\{\mathbf{M}_{s,i}\}(\mathbf{r}) + \frac{1}{2}\hat{\mathbf{n}}_i(\mathbf{r}) \times \mathbf{J}_{s,i}(\mathbf{r}) + j\frac{1}{2\eta_i k_i}[\nabla_s \cdot \mathbf{M}_{s,i}(\mathbf{r})]\hat{\mathbf{n}}_i(\mathbf{r}) = \mathbf{H}_i^{\text{inc}}(\mathbf{r}), \mathbf{r} \in \partial\Omega_i. \tag{39}$$

Equation (38) is the electric field integral equation (EFIE) and (39) the magnetic field integral equation (MFIE).

Formulations that use either the EFIE or the MFIE for the solution of a scattering problem often lead to internal resonances, which produce inaccurate results, especially at a resonance frequency of the scatterer [45]. To avoid this major problem, different combinations of the EFIE and MFIE have been proposed. Surface integral equation formulations, which are free of internal resonances, can thus be obtained by summing, with appropriate coefficients, the EFIEs and the MFIEs for all computational domains, which leads to a final matrix system of combined integral equations that are solved simultaneously [46]. A conventional approach to derive SIE formulations involves the tangential traces of the EFIE and MFIE representations, along with the boundary conditions. Thus, two categories of SIE formulations are produced. The N-Formulations are produced by combining the following tangential components

$$(\text{N}-\text{EFIE})_i : [\hat{\mathbf{n}}(\mathbf{r}) \times (\text{EFIE})]_{\partial\Omega_i}, \tag{40}$$

$$(\text{N}-\text{MFIE})_i : [\hat{\mathbf{n}}(\mathbf{r}) \times (\text{MFIE})]_{\partial\Omega_i}, \tag{41}$$

and the T-Formulations consist of

$$(\text{T}-\text{EFIE}) : [-\hat{\mathbf{n}}(\mathbf{r}) \times \hat{\mathbf{n}}(\mathbf{r}) \times (\text{EFIE})]_{\partial\Omega_i}, \tag{42}$$

$$(\text{T}-\text{MFIE}) : [-\hat{\mathbf{n}}(\mathbf{r}) \times \hat{\mathbf{n}}(\mathbf{r}) \times (\text{MFIE})]_{\partial\Omega_i}, \tag{43}$$

where $\hat{\mathbf{n}}(\mathbf{r})$ is the outward unit normal vector on the closed surface $\partial\Omega_i$.

C. N-Formulations

Consider a boundary $\partial\Omega$ between two adjacent domains Ω_ℓ and Ω_m. The selection $\hat{\mathbf{n}}(\mathbf{r}) = \hat{\mathbf{n}}_\ell(\mathbf{r})$ is made, where $\hat{\mathbf{n}}_\ell(\mathbf{r})$ is the normal unit vector of $\partial\Omega$ towards Ω_m. Then, N-EFIE and N-MFIE take the following forms in the domains Ω_ℓ and Ω_m:

$(\text{N}-\text{EFIE})_\ell$:

$$+\eta_\ell \mathscr{D}_\ell^n \{\mathbf{J}_{s,\ell}\}(\mathbf{r}) + \left(\mathscr{K}_\ell^n + \frac{\mathscr{I}}{2}\right)\{\mathbf{M}_{s,\ell}\}(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}_\ell^{\text{inc}}(\mathbf{r}), \tag{44}$$

$(\text{N}-\text{EFIE})_m$:

$$-\eta_m \mathscr{D}_m^n \{\mathbf{J}_{s,\ell}\}(\mathbf{r}) - \left(\mathscr{K}_m^n - \frac{\mathscr{I}}{2}\right)\{\mathbf{M}_{s,\ell}\}(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}_m^{\text{inc}}(\mathbf{r}), \tag{45}$$

$(\text{N}-\text{MFIE})_\ell$:

$$-\left(\mathscr{K}_\ell^n + \frac{\mathscr{I}}{2}\right)\{\mathbf{J}_{s,\ell}\}(\mathbf{r}) + \eta_\ell^{-1}\mathscr{D}_\ell^n \{\mathbf{M}_{s,\ell}\}(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}_\ell^{\text{inc}}(\mathbf{r}), \tag{46}$$

$(\text{N}-\text{MFIE})_m$:

$$+\left(\mathscr{K}_m^n - \frac{\mathscr{I}}{2}\right)\{\mathbf{J}_{s,\ell}\}(\mathbf{r}) - \eta_m^{-1}\mathscr{D}_m^n \{\mathbf{M}_{s,\ell}\}(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}_m^{\text{inc}}(\mathbf{r}), \tag{47}$$

where $\mathscr{I}$ is the identity operator and

$$\mathscr{D}_i^n \{\mathscr{X}\}(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathscr{D}_i \{\mathscr{X}\}(\mathbf{r}), \tag{48}$$

$$\mathscr{K}_i^n \{\mathscr{X}\}(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathscr{K}_i \{\mathscr{X}\}(\mathbf{r}). \tag{49}$$

Different N-Formulations can then be obtained by combining the previous equations with different coefficients, as shown below

$$m_\ell^N (\text{N}-\text{MFIE})_\ell + m_m^N (\text{N}-\text{MFIE})_m, \tag{50}$$

$$e_\ell^N (\text{N}-\text{EFIE})_\ell + e_m^N (\text{N}-\text{EFIE})_m. \tag{51}$$

The most popular formulation of this kind is mN-Müller [47] with coefficients

$$m_\ell^N = \frac{\mu_\ell}{\mu_m + \mu_\ell}, m_m^N = \frac{\mu_m}{\mu_m + \mu_\ell}, \tag{52}$$

and

$$e_\ell^N = \frac{\varepsilon_\ell}{\varepsilon_m + \varepsilon_\ell}, e_m^N = \frac{\varepsilon_m}{\varepsilon_m + \varepsilon_\ell}. \tag{53}$$

The N-Müller (with coefficients: $m_\ell^N = \mu_\ell, m_m^N = \mu_m$, $e_\ell^N = \varepsilon_\ell, e_m^N = \varepsilon_m$) [25] and mN-Müller [47] formulations have very fast rates of convergence when iterative solvers are used [48]. This happens because of the identity operator that appears on the diagonal of the system matrix. Thus, these formulations present a low condition number and fast convergence [48]. However, there are losses in terms of accuracy with the use of divergence-conforming basis functions of the lowest order, which will be discussed in section IVB.

D. T-Formulations

T-Formulations are a linear combination of the T-EFIE (42) and the T-MFIE (43). Again, consider a boundary $\partial\Omega$ between two adjacent domains Ω_ℓ and Ω_m. The selection $\hat{\mathbf{n}}(\mathbf{r}) = \hat{\mathbf{n}}_\ell(\mathbf{r})$ is made, where $\hat{\mathbf{n}}_\ell(\mathbf{r})$ is the normal unit vector of $\partial\Omega$ towards Ω_m. Then, T-EFIE and T-MFIE take the following forms in the domains Ω_ℓ and Ω_m:

$(\mathrm{T}-\mathrm{EFIE})_\ell$:

$$+\eta_\ell \mathscr{D}_\ell^t\left\{\mathbf{J}_{s,\ell}\right\}(\mathbf{r}) + \left(\mathscr{K}_\ell^t - \frac{\mathscr{I}_n}{2}\right)\left\{\mathbf{M}_{s,\ell}\right\}(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r})\times\hat{\mathbf{n}}(\mathbf{r})\times\mathbf{E}_\ell^{\mathrm{inc}}(\mathbf{r}), \quad (54)$$

$(\mathrm{T}-\mathrm{EFIE})_m$:

$$-\eta_m \mathscr{D}_m^t\left\{\mathbf{J}_{s,\ell}\right\}(\mathbf{r}) - \left(\mathscr{K}_m^t + \frac{\mathscr{I}_n}{2}\right)\left\{\mathbf{M}_{s,\ell}\right\}(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r})\times\hat{\mathbf{n}}(\mathbf{r})\times\mathbf{E}_m^{\mathrm{inc}}(\mathbf{r}), \quad (55)$$

$(\mathrm{T}-\mathrm{MFIE})_\ell$:

$$-\left(\mathscr{K}_\ell^t - \frac{\mathscr{I}_n}{2}\right)\left\{\mathbf{J}_{s,\ell}\right\}(\mathbf{r}) + \eta_\ell^{-1}\mathscr{D}_\ell^t\left\{\mathbf{M}_{s,\ell}\right\}(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r})\times\hat{\mathbf{n}}(\mathbf{r})\times\mathbf{H}_\ell^{\mathrm{inc}}(\mathbf{r}), \quad (56)$$

$(\mathrm{T}-\mathrm{MFIE})_m$:

$$+\left(\mathscr{K}_m^t + \frac{\mathscr{I}_n}{2}\right)\left\{\mathbf{J}_{s,\ell}\right\}(\mathbf{r}) - \eta_m^{-1}\mathscr{D}_m^t\left\{\mathbf{M}_{s,\ell}\right\}(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r})\times\hat{\mathbf{n}}(\mathbf{r})\times\mathbf{H}_m^{\mathrm{inc}}(\mathbf{r}), \quad (57)$$

where

$$\mathscr{D}_i^t\{\mathscr{X}\}(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r})\times\hat{\mathbf{n}}(\mathbf{r})\times\mathscr{D}_i\{\mathscr{X}\}(\mathbf{r}), \quad (58)$$

$$\mathscr{K}_i^t\{\mathscr{X}\}(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r})\times\hat{\mathbf{n}}(\mathbf{r})\times\mathscr{K}_i\{\mathscr{X}\}(\mathbf{r}), \quad (59)$$

$$\mathscr{I}_n = \hat{\mathbf{n}}\times\mathscr{I}. \quad (60)$$

Different T-Formulations can then be obtained by combining the previous equations with different coefficients, as shown below

$$e_\ell^T(\mathrm{T}-\mathrm{EFIE})_\ell + e_m^T(\mathrm{T}-\mathrm{EFIE})_m, \quad (61)$$

$$m_\ell^T(\mathrm{T}-\mathrm{MFIE})_\ell + m_m^T(\mathrm{T}-\mathrm{MFIE})_m. \quad (62)$$

The most popular formulation of this kind is T-PMCHWT [49] with coefficients

$$m_\ell^T = -m_m^T = e_\ell^T = -e_m^T = 1. \quad (63)$$

The T-PMCHWT [49], which is a Fredholm equation of the first kind, includes weakly singular integral operators. Hence, it presents very slow convergence compared to other N- and T-Formulations, because of the weak diagonal contributions in the system matrix [50]. However, its convergence can be improved by applying a preconditioning technique. The Calderon multiplicative preconditioner (CMP) has proved to be very effective in that context [51, 52].

E. Combined Field Integral Equations (CFIE)

The CFIE formulations are obtained by a linear combination of T-EFIE, N-EFIE, T-MFIE, and N-MFIE equations, in the adjacent regions:

$$e_m^N\eta_m^{-1}(\mathrm{N}-\mathrm{EFIE})_m + e_m^T\eta_m^{-1}(\mathrm{T}-\mathrm{EFIE})_m + m_m^N(\mathrm{N}-\mathrm{MFIE})_m + m_m^T(\mathrm{T}-\mathrm{MFIE})_m, \quad (64)$$

$$e_\ell^N(\mathrm{N}-\mathrm{EFIE})_\ell + e_\ell^T(\mathrm{T}-\mathrm{EFIE})_\ell + m_\ell^N\eta_\ell(\mathrm{N}-\mathrm{MFIE})_\ell + m_\ell^T\eta_\ell(\mathrm{T}-\mathrm{MFIE})_\ell. \quad (65)$$

However, every CFIE formulation does not lead to accurate solutions. Setting to zero one of the coefficients in the previous equations [45] gives the following four categories of formulations: TENE-TM ($m^N = 0$), TE-THNH ($e^N = 0$), TENE-NH ($m^T = 0$), and NE-THNH ($e^T = 0$). For every possible combination of $m^N, m^T, e^N, e^T \in \{-1, 1\}$ it has been shown that the CFIE formulations are free of resonances and present accurate solutions [53]. A different kind of CFIE formulation is JMCFIE [54], which consists of the following set of equations:

$$(\mathrm{JCFIE}): -\eta_m^{-1}(\mathrm{T}-\mathrm{EFIE})_m + \eta_\ell^{-1}(\mathrm{T}-\mathrm{EFIE})_\ell + (\mathrm{N}-\mathrm{MFIE})_m + (\mathrm{N}-\mathrm{MFIE})_\ell, \quad (66)$$

$$(\mathrm{MCFIE}): -\eta_m(\mathrm{T}-\mathrm{MFIE})_m + \eta_\ell(\mathrm{T}-\mathrm{MFIE})_\ell + (\mathrm{N}-\mathrm{EFIE})_m + (\mathrm{N}-\mathrm{EFIE})_\ell. \quad (67)$$

This formulation has been shown to be very robust, with fast convergence for iterative solvers, accurate far-field representation [48], and high efficiency for the simulation of composite objects with junctions [54].

IV. DISCRETIZATION

The discretization of the surface(s) of the scatterer(s) is an essential step to numerically solve an electromagnetic scattering problem with the various formulations presented above. Consider the simple problem of a single scattering body that lies in a homogeneous background medium. In this case, the first step is to discretize the body's surface by implementing a surface triangulation which will generate a mesh $\mathscr{M}$ with N_e edges.

To use the Method of Moments (MoM), both the surface electric and magnetic current densities must be expanded with basis functions to calculate them numerically on each edge of $\mathscr{M}$ in terms of their expansion coefficients:

$$\mathbf{J}_{s,i} = \sum_{n=1}^{N_e}\alpha_n\mathbf{f}_n(\mathbf{r}), \quad (68)$$

$$\mathbf{M}_{s,i} = \sum_{n=1}^{N_e}\beta_n\mathbf{f}_n(\mathbf{r}), \quad (69)$$

where $\mathbf{f}_n(\mathbf{r})$ are the basis functions. Also, a testing procedure with a testing function is required, to convert the integral equation into a matrix equation. The approach where the same function is used as basis and for testing is the well known Galerkin method [55]. The choice of

the proper basis and testing functions is essential for the implementation of the MoM, to obtain a final system that leads to accurate solutions. Hence, a mathematical analysis regarding function spaces and integral operators is needed to properly select the testing and basis functions.

A. Function spaces

Let us look at a scalar boundary value problem with a solution Ψ. Sobolev considered that inside any bounded medium, the energy should be finite, which leads to the requirement that both Ψ and $\nabla\Psi$ should be square integrable in a bounded domain Ω [56]. For interior problems this space is presented as $H^1(\Omega)$. Thus, the space of square integrable scalar functions in a bounded domain Ω is defined as $H^0(\Omega) = L^2(\Omega)$. If the domain Ω is unbounded, then there is a local definition of the square integrable feature in every bounded subset of Ω [56].

We are concerned about the boundary values on $\partial\Omega$, since the surface electric and magnetic current densities are on the boundary. With the help of the trace theorem it has been shown that the set of all boundary values in $H^1(\Omega)$ form the Hilbert space $H^{1/2}(\partial\Omega)$ [57], which is smaller than $L^2(\partial\Omega)$. Moreover, if $\nabla^2\Psi$ is also square integrable, all the normal derivatives of functions in $H^1(\Omega)$ form the space $H^{-1/2}(\partial\Omega)$, which is the dual of $H^{1/2}(\partial\Omega)$. Regarding the electromagnetic vector fields, they belong to $H^1(\Omega)$ given that every field component is in this space.

A curl Sobolev space is defined as a space in which functions and their curls are square integrable [58]

$$H(\mathrm{curl},\Omega) = \left\{\mathbf{f} : \mathbf{f} \in L^2(\Omega), \nabla\times\mathbf{f} \in L^2(\Omega)\right\}. \quad (70)$$

Poynting's theorem implies that energy is bounded if both the electric and magnetic field intensities are square integrable over any bounded subdomain of Ω. By considering Maxwell's equations, the energy is bounded if $\mathbf{E}$, $\mathbf{H}$, $\nabla\times\mathbf{E}$, $\nabla\times\mathbf{H}$ are square integrable in bounded subdomains, which means that both $\mathbf{E}$, $\mathbf{H} \in H(\mathrm{curl},\Omega)$. The analogous divergence Sobolev space is [58]

$$H(\mathrm{div},\Omega) = \left\{\mathbf{f} : \mathbf{f} \in L^2(\Omega), \nabla\cdot\mathbf{f} \in L^2(\Omega)\right\}, \quad (71)$$

which includes both the electric and magnetic flux densities: $\mathbf{D}$, $\mathbf{B} \in H(\mathrm{div},\Omega)$. In order to analyse SIEs, we need the trace spaces that are presented below, which illustrate the effect of applying the trace operators $(-\hat{\mathbf{n}}\times\hat{\mathbf{n}}\times)$ and $(\hat{\mathbf{n}}\times)$ to a function $\mathbf{f} \in H(\mathrm{curl},\Omega)$:

$$-\hat{\mathbf{n}}\times\hat{\mathbf{n}}\times\mathbf{f} : H(\mathrm{curl},\Omega) \longrightarrow H^{-1/2}(\mathrm{curl},\partial\Omega), \quad (72)$$

$$\hat{\mathbf{n}}\times\mathbf{f} : H(\mathrm{curl},\Omega) \longrightarrow H^{-1/2}(\mathrm{div},\partial\Omega), \quad (73)$$

where $\partial\Omega$ is the boundary of domain Ω and

$$H^{-1/2}(\mathrm{curl},\partial\Omega) = \left\{\mathbf{f} : \mathbf{f} \in H^{-1/2}(\partial\Omega), \hat{\mathbf{n}}\cdot\nabla_s\times\mathbf{f} \in H^{-1/2}(\partial\Omega)\right\}, \quad (74)$$

$$H^{-1/2}(\mathrm{div},\partial\Omega) = \left\{\mathbf{f} : \mathbf{f} \in H^{-1/2}(\partial\Omega), \nabla_s\cdot\mathbf{f} \in H^{-1/2}(\partial\Omega)\right\}, \quad (75)$$

where ∇_s is the surface gradient on the boundary. The space $H^{-1/2}(\mathrm{div},\partial\Omega)$, which includes both the surface electric current density ($\mathbf{J} = \hat{\mathbf{n}}\times\mathbf{H}$) and the surface magnetic current density ($\mathbf{M} = -\hat{\mathbf{n}}\times\mathbf{E}$), is the L^2 dual of $H^{-1/2}(\mathrm{curl},\partial\Omega)$ [59].

B. Basis functions

In general, the continuity equations for both the surface electric and magnetic current densities impose a physical requirement on the basis function, which should be able to represent properly the quantities $\nabla_s\cdot\mathbf{J_s}$ and $\nabla_s\cdot\mathbf{M_s}$, that are related to the surface electric ($\rho_{e,s}$) and magnetic ($\rho_{m,s}$) charge densities (multiplied with the factor $j\omega$). Thus, a good representation of both current densities requires the use of a divergence-conforming basis function. The mathematical approach of the previous section showed that the electric and magnetic current densities belong to $H^{-1/2}(\mathrm{div},\partial\Omega)$, which confirms that basis functions have to be divergence-conforming. The two most popular functions of this category are the Rao-Wilton-Glisson (RWG) and the quasi-curl-conforming Buffa-Christiansen (BC) functions that will be described next.

1. Rao-Wilton-Glisson (RWG)

The most common basis function is the RWG [60], which is the lowest order divergence-conforming function. Consider a triangular mesh $\mathscr{M}$ on a surface; a RWG function $\mathbf{f}_n(\mathbf{r})$ is defined in Fig. 3 for every pair of adjacent triangles T^+ and T^- with a common edge ℓ_n. The analytical expression for RWG is given by

$$\mathbf{f}_n(\mathbf{r}) = \begin{cases} +\dfrac{\ell_n}{2A^+}(\mathbf{r}-\mathbf{p}^+), & \mathbf{r}\in T^+ \\ -\dfrac{\ell_n}{2A^-}(\mathbf{r}-\mathbf{p}^-), & \mathbf{r}\in T^- \\ 0, & \text{otherwise} \end{cases}, \quad (76)$$

where $\mathbf{p}^+$ and $\mathbf{p}^-$ are the vertices of the two triangles T^+ and T^-, opposed to their common edge ℓ_n. Also, ℓ_n is the length of the common edge and A^+ and A^- are the areas of T^+ and T^-.

A main feature of RWG is that there is no normal component of the surface current density along the surrounding line boundary of the pair T^+, T^-, which means that line charges do not exist on it. Also, the component of the surface current density that is normal to the common edge ℓ_n is constant and continuous across ℓ_n. Furthermore, the surface charge density is constant in each

triangular element, since

$$\nabla_s \cdot \mathbf{f}_n(\mathbf{r}) = \begin{cases} +\dfrac{\ell_n}{A^+}, & \mathbf{r} \in T^+ \\ -\dfrac{\ell_n}{A^-}, & \mathbf{r} \in T^- \\ 0, & \text{otherwise} \end{cases}, \qquad (77)$$

with the total charge on each pair accumulating to zero.

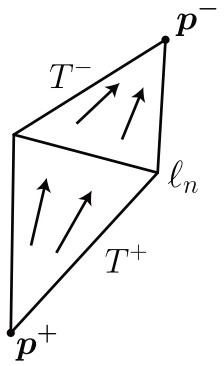


Fig. 3. RWG function on a pair of adjacent triangles.

2. Half-RWG

The half-RWG basis function is a modified version of its original counterpart and is defined only in a single triangle of mesh $\mathscr{M}$ [61], as shown in Fig. 4. Every edge $\ell_n \in \mathscr{M}$ has an arbitrarily generated direction. However, this direction is fixed, so the edge vector $\boldsymbol{\ell}_n$ of each edge is constant. Also, the reference normal unit vector $\hat{\mathbf{n}}^{ref}$ points towards the direction that is the result of the counterclockwise rotation of the triangle's vertices (their order is initially defined once for each triangle). The half-RWG function associated with an edge ℓ_n of a mesh triangle T with area A is defined as

$$\mathbf{h}_n(\mathbf{r}) = \begin{cases} \dfrac{\pm \ell_n}{2A}(\mathbf{r}-\mathbf{p}), & \mathbf{r} \in T \\ 0, & \text{otherwise} \end{cases}, \qquad (78)$$

where $\mathbf{p}$ is the vertex of T across ℓ_n. The sign of $\mathbf{h}_n(\mathbf{r})$ in each of the problem's domains Ω_i is different for opposite sides of the boundary between two adjacent domains. Hence, if Ω_ℓ and Ω_m have a common boundary, then $\mathbf{h}_{n,\ell}(\mathbf{r}) = -\mathbf{h}_{n,m}(\mathbf{r})$. This means that the boundary conditions regarding the surface electric and magnetic current densities are satisfied by half-RWG, when using the same expansion coefficients. As a result, some signs in the formulations that were previously presented will change with the use of this function. By employing the continuity of the original RWG-functions, the cumulative RWG function is defined as

$$\mathbf{f}_n(\mathbf{r}) = \sum_{\{k:\ell_k=\ell_n\}} \mathbf{h}_k(\mathbf{r}), \qquad (79)$$

where $\mathbf{f}_n(\mathbf{r})$ refers to an edge ℓ_n and $\mathbf{h}_k(\mathbf{r})$ refers to a half-RWG basis function that borders ℓ_n.

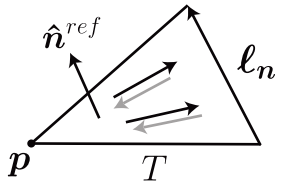


Fig. 4. Half-RWG function defined on a single mesh triangle.

3. Buffa-Christiansen (BC)

Another divergence-conforming and quasi-curl-conforming basis function is the BC function, which is defined on a barycentric refinement $\mathscr{M}_b$ of the original triangular mesh $\mathscr{M}$ [62]. Essentially, it is a linear combination of a set of RWG functions which are defined on $\mathscr{M}_b$. However, a BC function is associated with an edge of the original mesh $\mathscr{M}$.

Consider a reference edge ℓ_n on the original mesh $\mathscr{M}$. The barycentric refinement $\mathscr{M}_b$ is presented in Fig. 5. Around the right and left vertices of the reference edge there are $N_c = 4$ and $N_c^* = 5$ triangles, respectively, that belong to $\mathscr{M}$. In Fig. 5 the plus (+) and minus (−) signs show the appropriate direction of the numbered RWG functions (or half-RWG for open surfaces) on the new edges of $\mathscr{M}_b$. The linear combination of these functions, with appropriate signs and coefficients, synthesizes the BC basis function of the reference edge. The aforementioned coefficients are defined as follows [51],

$$c_i = \begin{cases} \dfrac{1}{2l_0}, & i = 0 \\ \dfrac{N_c - i}{2l_i N_c}, & i = 1, \ldots, 2N_c - 1 \end{cases}, \qquad (80)$$

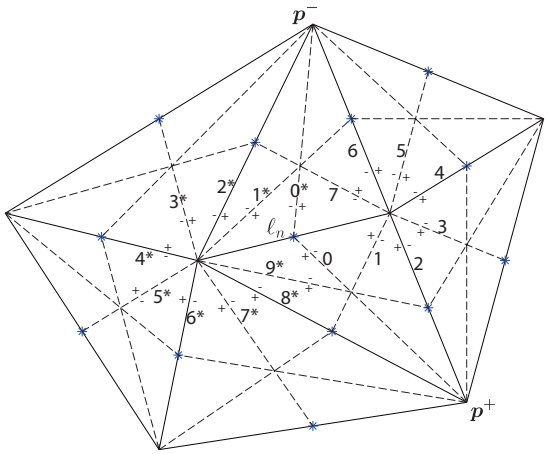


Fig. 5. Barycentically refined mesh $\mathscr{M}_b$.

$$c_{i^*} = \begin{cases} -\dfrac{1}{2l_{0^*}}, & i^* = 0 \\ -\dfrac{N_c^* - i^*}{2l_{i^*}N_c^*}, & i^* = 1^*, ..., 2N_c^* - 1\,. \end{cases} \tag{81}$$

The form of the BC for the reference edge ℓ_n is shown in Fig. 6.

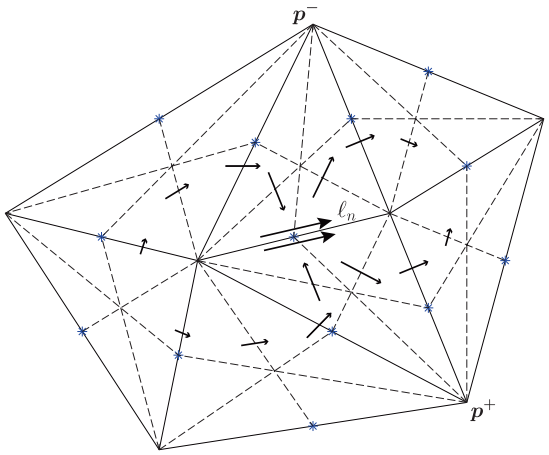


Fig. 6. Buffa-Christiansen basis function.

The BC functions present however a significant drawback regarding the representation of the surface charge density since they model the surface charge density as a constant function around the vertices of the original mesh $\mathscr{M}$ [42]. The RWG functions present constant surface charge density inside the triangular elements, as mentioned in the previous section. Hence, the latter are more appropriate for modeling discontinuous and singular surface charge density near sharp corners [63].

4. Trintinalia-Ling (TL)

The TL basis function [64], also called linear-linear (LL) basis function [65], can be perceived as a decomposition of the RWG basis function. Indeed, consider a triangular mesh $\mathscr{M}$ on a surface; the first and the second kind of LL functions are defined for every pair of adjacent triangles T^+ and T^- with a common edge ℓ_n, as shown below. Figures 7 and 8 show the form of both kinds of LL basis functions. The grayscale gradient inside the triangles illustrates the magnitude of the basis function. The analytical formulas of the first and second kinds of LL basis functions are given by (82) and (83), where ℓ_n is the length of the common edge and A^+ and A^- are the areas of T^+ and T^-.

For both kinds of LL functions, the spatial distribution is parallel to the surrounding edges where the magnitude of the basis function is nonzero (non-white color

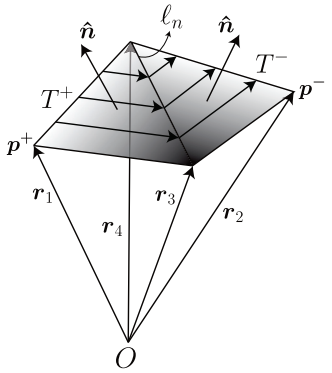


Fig. 7. First kind of LL basis function.

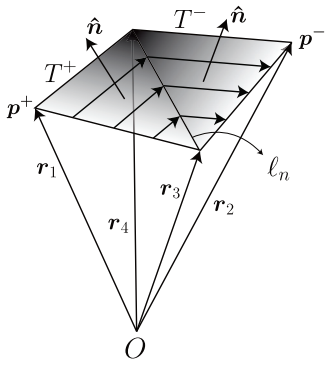


Fig. 8. Second kind of LL basis function.

$$\mathbf{f}_n^{LL,1}(\mathbf{r}) = \begin{cases} \dfrac{\ell_n}{4\left(A^+\right)^2}(\mathbf{r}-\mathbf{r}_1)\cdot\left[(\mathbf{r}_4-\mathbf{r}_1)\times\hat{\mathbf{n}}(\mathbf{r})\right](\mathbf{r}_3-\mathbf{r}_1), & \mathbf{r}\in T^+ \\ \dfrac{\ell_n}{4\left(A^-\right)^2}(\mathbf{r}-\mathbf{r}_2)\cdot\left[(\mathbf{r}_4-\mathbf{r}_2)\times\hat{\mathbf{n}}(\mathbf{r})\right](\mathbf{r}_3-\mathbf{r}_2), & \mathbf{r}\in T^- \\ 0, & \text{otherwise} \end{cases} \tag{82}$$

$$\mathbf{f}_n^{LL,2}(\mathbf{r}) = \begin{cases} \dfrac{\ell_n}{4\left(A^+\right)^2}(\mathbf{r}_1-\mathbf{r})\cdot\left[(\mathbf{r}_3-\mathbf{r}_1)\times\hat{\mathbf{n}}(\mathbf{r})\right](\mathbf{r}_4-\mathbf{r}_1), & \mathbf{r}\in T^+ \\ \dfrac{\ell_n}{4\left(A^-\right)^2}(\mathbf{r}_2-\mathbf{r})\cdot\left[(\mathbf{r}_3-\mathbf{r}_2)\times\hat{\mathbf{n}}(\mathbf{r})\right](\mathbf{r}_4-\mathbf{r}_2), & \mathbf{r}\in T^- \\ 0, & \text{otherwise} \end{cases} \tag{83}$$

in Figs. 7 and 8). Also, it varies linearly along those edges (maximum at the intersection with the reference edge ℓ_n). What is more, it is zero next to the surrounding edges where the magnitude of the basis function is zero (white color). Lastly, it exhibits a linear variation on the reference edge ℓ_n for both perpendicular and tangential directions.

As mentioned at the beginning of this section, LL functions are the decomposition of the RWG function, more specifically,

$$\mathbf{f}_n^{RWG}(\mathbf{r}) = \mathbf{f}_n^{LL,1}(\mathbf{r}) + \mathbf{f}_n^{LL,2}(\mathbf{r}) . \tag{84}$$

The surface divergence of these functions, which is proportional to the charge surface density, is expressed by

$$\nabla_s \cdot \mathbf{f}_n^{LL,1}(\mathbf{r}) = \nabla_s \cdot \mathbf{f}_n^{LL,2}(\mathbf{r}) = \begin{cases} +\dfrac{\ell_n}{2A^+}, & \mathbf{r} \in T^+ \\ -\dfrac{\ell_n}{2A^-}, & \mathbf{r} \in T^- \\ 0, & \text{otherwise} \end{cases} , \tag{85}$$

which is constant and the half of the surface divergence of the RWG function. From the last equation, we can see that the divergence of the LL functions is finite and comprehend that they are divergence-conforming.

C. Testing functions

As mentioned previously, a testing procedure with a function is required to convert the surface integral equation into a matrix equation. In this sense, the proper testing functions have to be selected for the MoM implementation, such that the final system leads to accurate solutions. In order to understand this, the mapping properties of surface integral operators have to be examined. Consider the following finite element spaces:

$$\mathscr{R}_{\mathscr{M}} = \left\{ \mathrm{RWG}_j \right\}_{j=1}^{N_e} , \tag{86}$$

$$\mathscr{R}^n_{\mathscr{M}} = \left\{ \hat{\mathbf{n}} \times \mathrm{RWG}_j \right\}_{j=1}^{N_e} , \tag{87}$$

$$\mathscr{B}_{\mathscr{M}} = \left\{ \mathrm{BC}_j \right\}_{j=1}^{N_e} , \tag{88}$$

$$\mathscr{B}^n_{\mathscr{M}} = \left\{ \hat{\mathbf{n}} \times \mathrm{BC}_j \right\}_{j=1}^{N_e} , \tag{89}$$

where $\mathscr{M}$ is the original mesh and N_e is the number of edges in $\mathscr{M}$. Regarding the previously presented formulations, the mapping properties of the discretized integral operators for RWG functions are [34]

$$\mathscr{D}^n \quad : \mathscr{R}_{\mathscr{M}} \longrightarrow \mathscr{B}_{\mathscr{M}} , \tag{90}$$

$$\mathscr{D}^t \quad : \mathscr{R}_{\mathscr{M}} \longrightarrow \mathscr{B}^n_{\mathscr{M}} , \tag{91}$$

$$\left(\mathscr{K}^n \pm \frac{\mathscr{I}}{2} \right) : \mathscr{R}_{\mathscr{M}} \longrightarrow \mathscr{R}_{\mathscr{M}} , \tag{92}$$

$$\left(\mathscr{K}^t \pm \frac{\mathscr{I}_n}{2} \right) : \mathscr{R}_{\mathscr{M}} \longrightarrow \mathscr{R}^n_{\mathscr{M}} . \tag{93}$$

As far as BC functions are concerned, the analogous mapping properties are the following [66]

$$\mathscr{D}^n \quad : \mathscr{B}_{\mathscr{M}} \longrightarrow \mathscr{R}_{\mathscr{M}} , \tag{94}$$

$$\mathscr{D}^t \quad : \mathscr{B}_{\mathscr{M}} \longrightarrow \mathscr{R}^n_{\mathscr{M}} , \tag{95}$$

$$\left(\mathscr{K}^n \pm \frac{\mathscr{I}}{2} \right) : \mathscr{B}_{\mathscr{M}} \longrightarrow \mathscr{B}_{\mathscr{M}} , \tag{96}$$

$$\left(\mathscr{K}^t \pm \frac{\mathscr{I}_n}{2} \right) : \mathscr{B}_{\mathscr{M}} \longrightarrow \mathscr{B}^n_{\mathscr{M}} . \tag{97}$$

Regarding the above spaces, the $\hat{\mathbf{n}} \times$ RWG function space is not an L^2 dual of the RWG space, but the $\hat{\mathbf{n}} \times$ BC function space is [34]. The same goes for the $\hat{\mathbf{n}} \times$ RWG and BC function spaces. In each formulation, the integral operators that provide the main contributions to the final matrix system (elements around the diagonal) have to be well tested. It has been shown that testing the surface integral operators with the dual of their range space leads to the most accurate results [66]. Thus, the basis function that is used to expand the main contributing integral operators will determine the choice of the testing function, as will be presented in section VB.

D. Singularity subtraction

The use of SIEs with the MoM produces singular integrals with weakly- or hyper-singular surface integral operators. These singularities appear when the basis and testing functions belong to the same triangular element or to adjacent triangles that share an edge or a vertex. The reason for these singularities is the denominator of Green's function, that goes to zero as $R = |\mathbf{r} - \mathbf{r}'| = 0$, when $\mathbf{r} = \mathbf{r}'$. The solution to this problem can be given by the application of a singularity subtraction method [67].

In order to implement the method, the Green's function in a domain Ω_i has to be expanded in a Taylor series, as follows,

$$G_i(\mathbf{r}, \mathbf{r}') = \frac{1}{4\pi} \left(\frac{1}{R} + jk_i - \frac{k_i^2 R}{2} + \frac{jk_i^2 R^2}{6} - ... \right) , \tag{98}$$

where the odd terms are the singular ones ($q = -1, 1, 3, ...$). In this sense, Green's function can be divided into a free of singularities smooth part and a singular part, hence

$$G_i(\mathbf{r}, \mathbf{r}') = G_i^s(\mathbf{r}, \mathbf{r}') + T_i(\mathbf{r}, \mathbf{r}') , \tag{99}$$

where

$$G_i^s(\mathbf{r}, \mathbf{r}') = \frac{e^{jk_i R}}{4\pi R} - \frac{1}{4\pi R} + \frac{k_i^2 R}{8\pi} , \tag{100}$$

$$T_i(\mathbf{r}, \mathbf{r}') = \frac{1}{4\pi R} - \frac{k_i^2 R}{8\pi} . \tag{101}$$

Thus, the smooth part $G_i^s(\mathbf{r}, \mathbf{r}')$ can be accurately integrated numerically and the singular part $T_i(\mathbf{r}, \mathbf{r}')$ can be integrated semi-analytically with the help of closed-form relations. In Fig. 9 we present the normalized values of the real part of the scalar Green's function, the real part of the smoothed term $G_i^s(\mathbf{r}, \mathbf{r}')$ and the singular term $T_i(\mathbf{r}, \mathbf{r}')$ as a function of the electrical distance $k_i R$.

The subtraction of the first odd term would be enough for making the smooth part non-singular, but in this case Green's function would have a discontinuous derivative at $R = 0$. Also, the term $k_i R$ is small for the case of singularities so the first and the second odd terms are enough for the definition of $G_i^s(\mathbf{r}, \mathbf{r}')$, since odd terms of higher orders diminish rapidly for small values of $k_i R$ [68].

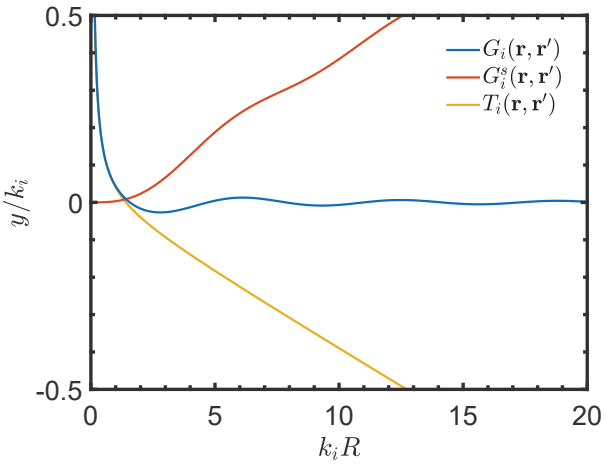


Fig. 9. Graphical representation of singularity subtraction.

The integrals that occur from the implementation of the singularity subtraction method are the following

$$I_1^q = \int_{\partial\Omega_i} \mathrm{d}S \left[\nabla \cdot \mathbf{f}_m(\mathbf{r})\right] \int_{\partial\Omega_i} \mathrm{d}S' R^q \nabla' \cdot \mathbf{f}_n(\mathbf{r}') \,, \quad (102)$$

$$I_2^q = \int_{\partial\Omega_i} \mathrm{d}S \mathbf{f}_m(\mathbf{r}) \cdot \int_{\partial\Omega_i} \mathrm{d}S' R^q \mathbf{f}_n(\mathbf{r}') \,, \quad (103)$$

$$I_3^q = \int_{\partial\Omega_i} \mathrm{d}S \mathbf{f}_m(\mathbf{r}) \cdot \int_{\partial\Omega_i} \mathrm{d}S' \left(\nabla' R^q\right) \times \mathbf{f}_n(\mathbf{r}') \,, \quad (104)$$

where $\mathbf{f}_m(\mathbf{r})$ is the testing function, $\mathbf{f}_n(\mathbf{r})$ is the basis function and $q = -1, 1$. The aforementioned integrals are solved semi-analytically, which means that the inner integrals are solved analytically for arbitrary $\mathbf{r}$, with closed-form expressions, and the outer integrals are solved numerically. The closed-form relations for the inner integrals can be found in [68].

V. FORMULATION AND DISCRETIZATION COMPARISONS

The previous considerations highlight the need to discuss in the following sections many different aspects of SIE methods, including the comparison between different formulations, as well as different basis and testing functions.

A. Comparison of different formulations

In this section, we compare the different SIE formulations – namely mN-Müller, T-PMCHWT and JMCFIE – for the case of a spherical gold nanostructure with a radius $r = 100$ nm in an air background, which presents the advantage that the analytical Mie solution can be considered as reference. A detailed discussion of the results for this case, as well as for some more general shape scatterers, can be found in [70]. We assume that the incident plane wave is propagating along the z-axis and has linear polarization towards the x-axis. The wavelength is $\lambda = 550$ nm (monochromatic results, such as the bistatic scattering cross section in Fig. 10) and the following figures were obtained with RWG basis functions and the Galerkin method.

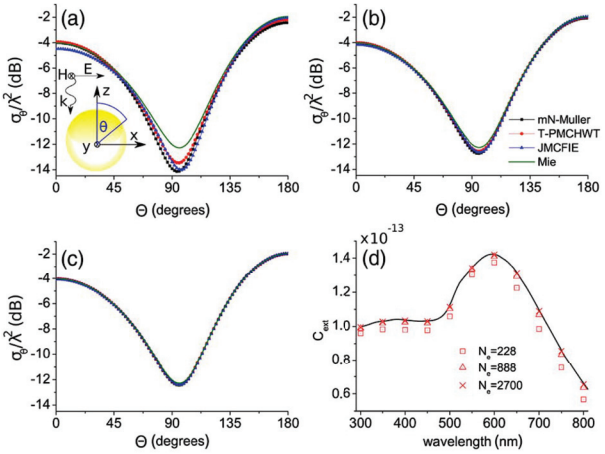


Fig. 10. Bistatic scattering cross section of a gold sphere ($r = 100$ nm) for three mesh densities with (a) $N_e = 228$, (b) $N_e = 888$, (c) $N_e = 2700$ number of edges and (d) extinction C_{ext} spectrum obtained by the T-PMCHWT for the three meshes compared to the Mie solution. Adapted with permission from [70] © Optica Publishing Group.

It is expected that, as the number of unknowns increases, the different formulations tend to approximate better the Mie solution. However, T-PMCHWT seems to be closer to the Mie solution than other formulations for a smaller number of unknowns, see Fig. 10 (a). The very good matching of this formulation with the analytical solution is explained in Fig. 10 (d), where the spectrum of the sphere's extinction cross section C_{ext}, which is extracted via T-PMCHWT, is compared to the Mie theory solution.

As shown in Fig. 11 (b), all three formulations present the same level of accuracy in the near-field zone, where the tangential component of the scattered electric field intensity (on the sphere) is considered for the error extraction. However, in the far-zone, where C_{ext} is taken into account for error calculations, JMCFIE performs the best, followed by T-PMCHWT. The mN-Müller presents the highest error of the three, see Fig. 11 (a). In [73] the strong material dependencies of conventional

Table 1: Comparison of different formulations

Set of Formulations	Test Function	Basis Function	Reference
T-PMCHWT, CTF, ICTF	RWG	RWG	[69]
T-PMCHWT, mN-Müller, JMCFIE, NFM, DDA	RWG	RWG	[70]
T-PMCHWT, N-Müller, CTF, CNF, JMCFIE	RWG	RWG	[71]
T-PMCHWT, CTF, CNF, N-Müller, mN-Müller, JMCFIE	RWG	RWG	[48]
T-PMCHWT, CTF, CNF, N-Müller, mN-Müller, JMCFIE	TL	TL	[48]
T-PMCHWT, N-Müller, T-Müller	RWG	RWG	[47]
T-PMCHWT, CTF, CNF, N-Müller, T-PMCHWT (CMP)	RWG	RWG	[72]

formulations, such as the normal and scaled forms of T-PMCHWT, CTF, JMCFIE and others, are examined and their performances for different plasmonic nanostructure problems are presented. In Table 1 we present some of the papers that examine and compare different formulations of the SIE method, including some that were not discussed before; the combined tangential formulation (CTF), the improved combined tangential formulation (ICTF), the null field method (NFM) and the combined normal formulation (CNF), while CMP in the last row of Table 1 stands for Calderon multiplicative preconditioner [51, 52].

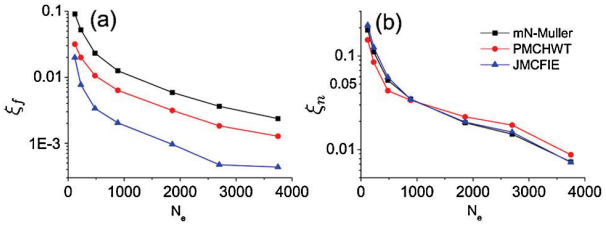


Fig. 11. Error in (a) the far zone ξ_f and (b) the near zone ξ_n of a gold sphere ($r = 100$ nm) obtained with the mN-Müller, T-PMCHWT, and JMCFIE formulations as a function of the number of mesh edges N_e. Adapted with permission from [70] © Optica Publishing Group.

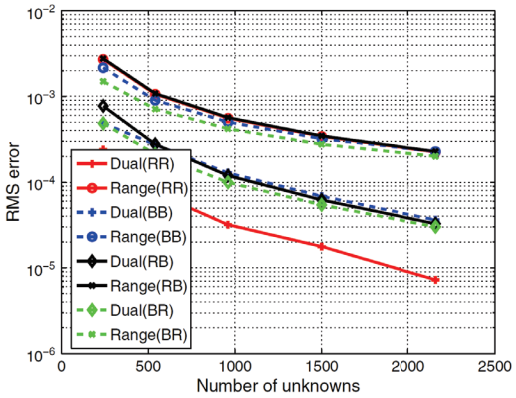


Fig. 12. RMS error of the bistatic RCS in the E-plane computed with the T-PMCHWT formulation. Dielectric sphere with $k_0 a = 1$ and $\varepsilon_r = 4$, $\mu_r = 1$. Adapted with permission from [66] © John Wiley & Sons.

Table 2: T-PMCHWT test & basis function pairs

Method	J	M	EFIE	MFIE
Dual (RR)	RWG	RWG	RWG	RWG
Dual (BB)	BC	BC	BC	BC
Dual (RB)	RWG	BC	RWG	BC
Dual (BR)	BC	RWG	BC	RWG

B. Basis and testing functions comparisons

As mentioned in section IVC, the use of the dual of the range of integral operators for testing leads to more accurate results. For the integral equations of the first kind, like EFIE and T-PMCHWT, this is identical with Galerkin's method, but for the integral equations of the second kind, like MFIE and mN-Müller, this leads the use of the Petrov–Galerkin method with appropriate curl-conforming testing functions. Moreover, the use of hybrid meshes introduces additional complexity and intriguing capabilities, while it also affects the condition number of the final system and the total simulation time, as shown in [74]. In [66] there is a very detailed examination of many different formulations, discretization schemes and problems, but we will focus mostly on the cases of T-PMCHWT and N-Müller.

In Fig. 12, the fact that the dual of the range of integral operators is better for testing in terms of accuracy is confirmed, while Table 2 presents all the different cases of dual testing for the T-PMCHWT formulation. Figure 12 shows that, for the T-PMCHWT formulation, when both RWG and BC are used as basis functions in a mixed discretization scheme (i.e. the cases Dual (RB) and Dual (BR) in the last two rows of Table 2), the results are worse than the Galerkin method with RWG functions as both testing and basis functions.

For the mN-Müller formulation, the reference Table 3 presents all the different cases of dual testing. Figure 13 presents similar results, but for the case of a mN-Müller formulation. In the case of the sphere, which is a smooth

Table 3: mN-Müller test & basis function pairs

Method	J	M	MFIE	EFIE
Dual (RR)	RWG	RWG	n×BC	n×BC
Dual (BB)	BC	BC	n×RWG	n×RWG
Dual (RB)	RWG	BC	n×BC	n×RWG
Dual (BR)	BC	RWG	n×RWG	n×BC

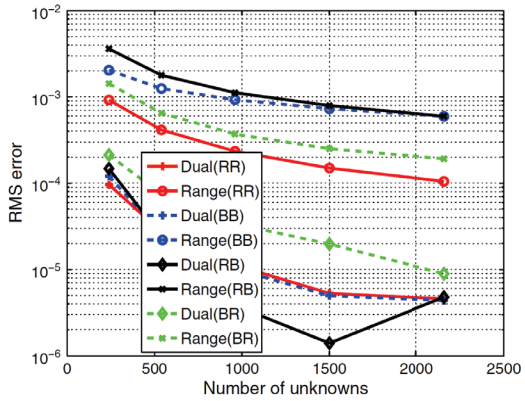


Fig. 13. RMS error of the bistatic RCS in the E-plane computed with the mN-Müller formulation. Dielectric sphere with $k_0a = 1$ and $\varepsilon_r = 4$, $\mu_r = 1$. Adapted with permission from [66] © John Wiley & Sons.

object, the mN-Müller gives a more accurate solution, compared to the results of the T-PMCHWT formulation, when the Dual(RB) mixed discretization scheme is used. However, for objects that are not smooth, such as cubes and prisms, the T-PMCHWT formulation is more accurate than the mN-Müller formulation. Lastly, a comprehensive comparison of different discretization schemes, including the various test and basis function pairings used in common formulations, is presented in Table 4.

VI. NUMERICAL SOLVERS AND APPROACHES

As mentioned in previous sections, integral equations are converted into the final linear system with the aid of a set of testing and basis functions. However, a significant challenge in SIE methods has been the appearance of large dense matrices after the discretization of complex real-life electromagnetic scattering problems. In the 1980s, the capabilities of MoM solvers were illustrated in [80], which highlighted that the largest problem solvable via direct lower–upper (LU) factorization within an hour was still relatively modest in size. The general consensus at the time was that MoM would not be a viable approach for real-world scattering scenarios due to its high computational cost and limited scalability, since it scaled as $\mathcal{O}(N^2)$ in terms of memory consumption while having a complexity of $\mathcal{O}(N^3)$ for direct solvers. In terms of iterative solvers, the dense matrix equations involving $\mathcal{O}(N)$ unknowns can be addressed via a Krylov-subspace algorithm, such as the generalized minimum residual (GMRES), conjugate gradient (CG), and bi-conjugate stabilized (BiCGStab) [81], but the aforementioned bottlenecks make their application to large-scale models impractical. Another key challenge of the SIE method is that some formulations (e.g. the EFIE and PMCHWT) belong to the first kind of integral equations. As a result, the final linear systems tend to be ill-conditioned, prompting the need for additional techniques, such as regularization or specialized preconditioning, to ensure that the resulting matrices are well-conditioned and can be solved efficiently. Also,

Table 4: Comparison of different discretization schemes

Formulations	Test & Basis Functions Pairs	References
MFIE	(RWG, RWG) and ($\hat{\mathbf{n}}$×RWG, $\hat{\mathbf{n}}$×RWG)	[75]
MFIE	(RWG, RWG), ($\hat{\mathbf{n}}$×RWG, $\hat{\mathbf{n}}$×RWG), and (monopolar RWG, monopolar RWG)	[76, 77]
MFIE	(LL, LL) and (RWG, RWG)	[65, 78]
PMCHWT	($\hat{\mathbf{n}}$×RWG, RWG)	[41]
Müller	(RWG, RWG) and ($\hat{\mathbf{n}}$×BC, RWG)	[41]
N-Müller	(RWG, RWG) and ($\hat{\mathbf{n}}$×BC, RWG)	[32]
EFIE	(RWG, RWG) and (BC, BC)	[66]
MFIE	($\hat{\mathbf{n}}$×BC, RWG) and ($\hat{\mathbf{n}}$×RWG, BC)	[66]
PMCHWT	(RWG, RWG), (BC, BC), (Mixed RWG-BC, Mixed RWG-BC), and (Mixed BC-RWG, Mixed BC-RWG)	[66]
mN-Müller	($\hat{\mathbf{n}}$×BC, RWG), (Mixed $\hat{\mathbf{n}}$×BC-$\hat{\mathbf{n}}$×RWG, Mixed RWG-BC), ($\hat{\mathbf{n}}$×RWG, BC), and (Mixed $\hat{\mathbf{n}}$×RWG-$\hat{\mathbf{n}}$×BC, Mixed BC-RWG)	[66]
MFIE, CFIE	Higher order test & basis functions	[79]
mN-Müller, JMCFIE	Higher order test & basis functions	[79]

an ill-conditioned final matrix system can occur when SIEs are used for low frequency simulations. However, the advancement of the Calderon preconditioner, along with the development of various numerical approaches for both direct [82] and iterative solvers [83, 84], and particularly the implementation of fast solvers, have helped mitigate these difficulties.

A. Low frequency breakdown

When the SIE method is used for electrically small yet complex geometries or with very dense discretizations, a recurring challenge known as the low frequency breakdown problem arises. This issue has been extensively analyzed in [85, 86]. One common strategy for handling the low-frequency regime, while preserving the correct physical behavior, is to employ loop-tree or loop-star basis functions combined with a suitable frequency normalization approach. Nevertheless, fast methods like the multi-level fast multipole algorithm (MLFMA) typically fail when the frequency is very low [87].

Even with loop-tree or loop-star basis functions, iterative solvers may still experience high iteration counts. The root cause is linked to the divergence properties of the RWG basis, which are not ideal for accurately representing charges at low frequencies [88]. To address this, a basis-rearrangement scheme was introduced in [88, 89]. Reshaping the way the basis functions are assembled substantially enhances the eigenvalue spectrum of the final system matrix and reduces notably the iteration count.

B. Calderon preconditioner

Calderon preconditioning is a regularization technique for SIEs [40, 51, 52, 90]. This approach utilizes the self-regularization property of the electric-field integral operator (EFIO), commonly referred to as the Calderon integral identity [56], which can be written as:

$$(\mathscr{D}^n)^2 = -\frac{1}{4}\mathscr{I} + (\mathscr{K}^n)^2\,, \tag{105}$$

where the notation of the previous sections is followed for the integral operators. The subscript i is missing, since for the final system to be formed, a summation over all domains has been executed. This identity demonstrates that when the EFIO ($\mathscr{D}^n$) is multiplied by itself, the resulting operator is better conditioned, particularly on sufficiently smooth surfaces where it approximates the summation of a compact and an identity operator. When the CMP (which was called Calderon multiplicative preconditioner in previous sections) is applied with the RWG and Buffa-Christiansen functions, it not only mitigates the ill-conditioning of the EFIO, but also maintains the solution accuracy [51].

Mathematically, the operator $(\mathscr{D}^n)^2$ establishes the mapping:

$$(\mathscr{D}^n)^2 : H^{-1/2}(\mathrm{div}, \partial\Omega) \to H^{-1/2}(\mathrm{div}, \partial\Omega)\,, \tag{106}$$

hence the application of the Calderon multiplicative preconditioner on the EFIE (CMP-EFIE) leads to a range space that is identical with that of the MFIE. This equivalence allows them to be effectively combined into a CFIE, which is inherently better conditioned than its non-preconditioned counterpart. However, the resulting CFIE system is not a resonance-free formulation, as the CMP-EFIE and the MFIE share the same resonances [91]. Lastly, Calderon preconditioning has also been extended for the PMCHWT formulation in the case of homogeneous isotropic [92] and chiral objects [93].

C. Iterative fast solvers

Iterative solutions require matrix-vector multiplications (MVMs), thus the development of a fast algorithm to solve the MoM equation requires the combination of an iterative method with a fast approach to compute the MVMs. Over the last few decades, a variety of fast solvers have been developed to overcome the high computational and memory costs traditionally associated with the MoM solution of SIEs. In general, fast algorithms broadly fall into two categories: kernel-dependent, which are influenced by the specific properties of the underlying Green's function, and kernel-independent, which leverage low-rank representations of system sub-blocks without requiring explicit knowledge of the integral operator's closed-form expansions.

One of the most impactful kernel-dependent solvers is the multilevel fast multipole method (MLFMM) [4, 94], which is an extension of the fast multipole method (FMM) [95, 96]. As presented in [97], in order to implement the MLFMM, the entire scatterer is initially enclosed within a large auxiliary cubic box, which is then divided into eight smaller cubes. This domain subdivision process continues recursively until the smallest cubes have an edge length that is comparable to the wavelength ($\approx \lambda/2$). Each cube at every level of this process is assigned an index. At the finest level, the cube containing each basis function is identified by comparing the coordinates of the basis function's center with the cube's center. Additionally, nonempty cubes are identified through a sorting process, and only these cubes are stored using a tree-structured data format [97]. Hence, the method organizes the interactions of basis-testing functions into a tree structure with multiple levels of hierarchically defined groups (clustering) of varying fineness, by utilizing the analytical expansion of Green's function and addition theorem [98]. Short-range (near-field) interactions are calculated directly and stored in memory, while long-range (far-field) interactions are performed efficiently using the factorization/diagonalization of the Green's function, thus reducing both the memory requirement and computational time complexity (MVM speed) to approximately

$\mathcal{O}(N\log N)$ for 3D electromagnetic scattering problems [98]. The introduction of this method marked a significant shift in computational electromagnetics, allowing MoM to tackle very large problems in terms of electrical size, while retaining its accuracy. The approach is widely parallelized in practice [99–101], making it one of the leading techniques for real-world large-scale modeling.

Another group of kernel-dependent solvers consists of the grid- or fast Fourier transform (FFT)-based techniques. The latter leverage the translation invariance of the kernel function, which can reduce the memory requirement and central processing unit (CPU) time complexity of 3D problems. Two of the most popular methods of this group are the pre-corrected FFT (p-FFT) and adaptive integral method (AIM), which typically achieve $\mathcal{O}(N^{1.5})$ for storage and $\mathcal{O}(N^{1.5}\log N)$ for MVM. More specifically, both methods rely on an equivalent source approximation. As described in [102], the acceleration of computations with the application of FFT requires the entire scatterer to be enclosed within an auxiliary rectangular domain. Afterwards, this auxiliary domain is recursively divided into a uniform Cartesian grid, ensuring that each small cube contains at most a few discretization elements. To perform the MVM with FFT, the original basis functions are mapped onto the Cartesian grid, which is achieved through basis transformation. Additionally, a fast high-order algorithm for solving surface scattering problems, utilizing a two-face equivalent source approximation, is presented in [103]. This method achieves computational complexities ranging from $\mathcal{O}(N^{6/5}\log N)$ to $\mathcal{O}(N^{4/3}\log N)$, by strategically positioning equivalent sources solely on the faces of cubic cells. In [104], the Sparse-Matrix/Canonical-Grid (SM/CG) technique is presented. Unlike the previously discussed approaches, it does not make use of equivalent sources; instead, it applies a Taylor series expansion of the Green's function on a regularly spaced canonical grid. The system matrix is then handled via an FFT-driven iterative scheme, where the number of Taylor expansion terms dictates the procedure. However, the SM/CG method demands detailed knowledge of the integral kernel, while its memory complexity scales as $\mathcal{O}(N^{1.5})$. In [105], the Quadrature Sampled Pre-Corrected Fast-Fourier Transform (QS-PCFFT) method was introduced. This approach projects the unknown currents onto a uniform grid and computes the discrete Fourier transform of the current directly by means of a discontinuous FFT, employing quadrature sampling for the currents. As a result, it achieves adjustable accuracy and exponential convergence. Furthermore, the Green's function interpolation combined with the FFT algorithm (GIFFT) [106] was developed to handle arrays of arbitrary shape. In GIFFT, an array mask function identifies array boundaries and determines where the Green's function should be interpolated. The MVMs in the iterative solver are then accelerated via FFT, reducing both storage requirements and overall solution time. Nonetheless, for SIEs, none of the aforementioned FFT-based techniques achieves asymptotically better performance than MLFMM. Finally, most iterative methods are sensitive to the condition number of matrix systems, and the number of iterations needed to reach the desired accuracy varies depending on the problem. Even though preconditioning techniques and domain decomposition methods have been developed to address many of these challenges, convergence remains not entirely predictable.

D. Direct fast solvers

In parallel, direct fast solvers utilizing hierarchical matrices (H-matrices) and matrix compression techniques have introduced kernel-independent algebraic methods. The latter rely on the observation that well-separated sub-blocks of the dense system matrix exhibit numerical rank deficiency. Algorithms such as the adaptive cross approximation (ACA) [107] compress these sub-blocks without requiring explicit expansions of the integral kernel.

The research of fast direct solvers for SIEs seek to address the drawbacks highlighted earlier. There are a number of fundamental studies that have laid the groundwork for direct solution techniques [108–111], many of which have since influenced advances in computational electromagnetics.

As presented in [82], early investigations include the IES^3 algorithm [112], direct treatments for 2D slender scatterers [113], and a compressed block decomposition (CBD) algorithm for capacitance extraction that combines matrix decomposition with singular value decomposition (SVD) [114]. Furthermore, an application of a multilevel nonuniform grid (MLNG) approach for direct inversion of the EFIE was presented some years ago in [115].

Meanwhile, the concept of H-matrices [116, 117], which systematically apply low-rank compression to well-separated matrix subgroups, has become a recognized framework for hierarchical direct solvers. Since this approach relies on certain smoothness criteria for the integral kernel, it typically applies to mid- and low-frequency regimes in electromagnetics, where the oscillatory nature of the kernel is less pronounced [118]. However, numerous implementations have adopted various H-matrix techniques [119–121]. Additional comparisons between multilevel block inversion and multilevel LU decomposition have also been made [122]. Another line of research incorporates the butterfly algorithm with randomized compression, achieving $\mathcal{O}(N\log^2 N)$ scaling and supporting large-scale 3D SIE simulations for PEC [123]. For penetrable scatterers, a quasi-block-Cholesky

(QBC) approach that investigates the checkerboard symmetry pattern in the PMCHWT system matrix was proposed in [124], while a butterfly-based multilevel matrix decomposition algorithm (MLMDA) for homogeneous penetrable objects was presented in [125]. Finally, in [82] a more detailed discussion on works that are based on the following matrix structures is provided: hierarchical off-diagonal low-rank (HODLR), hierarchically semiseparable (HSS), and H^2 matrices.

E. Higher order methods

Another development in integral equation methods has been the use of higher order isoparametric elements [126, 127]. Recently, it was shown that hybrid discretization schemes can affect the condition number of the final matrix system and the efficiency of the solution procedure [74]. This insight further reinforces the idea of using higher order techniques to improve the numerical behavior of SIEs. In many electromagnetic radiation and scattering problems, the primary source of numerical error often arises from inaccuracies in geometric modeling [87]. The use of these elements not only mitigates the aforementioned error, but also enables exponential error convergence in regions where the numerical solutions remain smooth [87].

There are two main approaches to employ higher order isoparametric elements in solving SIEs, namely the Galerkin [126, 127] and Nystrom methods [92, 105]. Both approaches have demonstrated higher order convergence rates for electromagnetic problems characterized by smooth solutions. A very detailed and comprehensive review of the higher order computational electromagnetics for antenna, wireless, and microwave engineering applications is presented in [128].

F. Parallelization and high-performance computing

High-performance computing (HPC) resources have been widely and effectively employed by researchers and engineers in the area of computational electromagnetics. Over time, developments in computing hardware and software have created extraordinary possibilities (and new challenges) for expanding the modeling techniques used for scattering problems. Crucially, these advances have taken place in parallel with progress in electromagnetic formulations and numerical algorithms. The research community has been especially resourceful in capitalizing on every available advance in hardware and technology.

Traditionally, two primary parallel programming paradigms, open multiprocessing (OpenMP) and the message passing interface (MPI), have dominated the landscape. These frameworks can facilitate massively parallel MLFMM simulations involving billions of unknowns [100, 129].

Table 5: Speed-up of bistatic radar cross section calculation of an aerocraft with CFIE at 1.5 GHz. Adapted with permission from [99] © IEEE.

	CPU (sec.)	GPU (sec.)	Speed-up
$\mathbf{V}_s$ and $\mathbf{V}_f$	52	19	2.8
$\mathbf{T}$	44	1	44.0
$\mathbf{Z}_{near}$	3735	30	124.5
BiCGStab	1911	653	2.9
Total time	5742	705	8.1

A historical perspective highlights the scale of these achievements. In 1988, Miller illustrated how the largest feasible problem within a one-hour time frame advanced from only around 100 unknowns in the early 1950s to approximately 6,000 unknowns (roughly a factor of 60) by the mid-1980s [80]. However, over the next few decades, this capability surged to approximately 10.5 billion unknowns within about 6.25 hours [100], yielding an increase on the order of 1.75 million. Thus, while the first half-century of computational electromagnetics and SIEs may have appeared to progress relatively slowly, the subsequent 30 years saw an explosive expansion of modeling capacity. Additionally, graphics processing units (GPUs), which were originally designed for computer graphics, evolved into indispensable hardware accelerators that can be used with the SIE method. As shown in Table 5, the utilization of GPUs can significantly enhance the efficiency of the SIE method [99]. More specifically, Table 5 includes time data regarding the calculation of radiation patterns of the basis functions $\mathbf{V}_s$, receiving patterns of the testing functions $\mathbf{V}_f$, the translator factor $\mathbf{T}$, and the assembly of the near-field system matrix $\mathbf{Z}_{near}$. Additional details about these quantities can be found in [99].

VII. CONCLUSION AND OUTLOOK

The motivation for this review article was to collect and present a complete and comprehensive overview of the SIE method theory, formulations, discretization schemes, and final matrix system considerations. This frequency domain computational electromagnetic technique, with its surface meshing approach, significantly decreases the number of unknowns for a given scattering problem and improves the analysis accuracy and efficiency. Firstly, by focusing on SIE's theoretical aspects, this paper provides a unified and systematic presentation of the electromagnetic analysis and the derivation of different popular formulations. Furthermore, a mathematical overview of function spaces is included, in order to highlight the importance of correctly combining basis and testing functions. Moreover, this work examines different aspects of a discretization scheme, such

as subtracting the Green's function singularity, which is essential for near-field representations, and applying different testing methods, in order to investigate procedures that are different from the Galerkin method. Hence, by delving into all the aforementioned parts of a complete SIE computational implementation, we emphasize that a surface integral equation formulation and the discretization process must lead to a unique solution and a conforming discretization procedure, i.e. the integral operators that are the main contributors regarding the off-diagonal elements of the final system matrix should be tested with the dual of their range space, since this leads to the most accurate results. In this manner, we review some of the most used SIE methods and illustrate whether they meet these conditions and how they perform compared to each other in terms of accuracy. Also, different discretization techniques using different basis and testing functions are investigated in terms of performance. Finally, we discuss the various challenges encountered in solving the final matrix equations, along with numerical strategies designed to mitigate their adverse computational impact. These strategies include the utilization of the Calderon preconditioner, the integration of advanced numerical and computational techniques for both direct and iterative solvers, the implementation of fast solvers, the application of higher order methods, and the effective utilization of HPC resources.

Despite all the significant work that has been done over the last decades, the method still faces some challenges, e.g. the dense mesh breakdown, the low frequency breakdown, the simulation of composite metallic and dielectric structures with junctions or objects of high conductivity that are modeled with impedance boundary conditions (IBCs).

Finally, we highlight some interesting research directions for the future in electromagnetic SIE methods, such as hybrid meshes [74] (e.g. the combination of triangular and quadrilateral elements), higher order and conformal elements, generalized MoM approaches [130], discontinuous Galerkin SIE methods [131], and source formulations [132]. Furthermore, one of the most significant challenges regarding the SIE method is to leverage hybrid parallelization paradigms and methodologies (e.g. combining MPI, OpenMP, and CUDA) to fully exploit the ever-growing variety and capability of modern HPC hardware [101]. In this context, the path to a general and fully optimized solution for SIE problems likely lies in combining multiple key elements – such as robust preconditioners, advanced fast and direct solvers, higher order conformal discretizations, and efficient parallelization strategies – into a synergistic framework that will be able to span across the entirety of the frequency and material spectra. By uniting these concepts, researchers and practitioners stand poised to overcome long-standing challenges and further advance the scope and efficiency of SIE-based electromagnetic simulations.

ACKNOWLEDGMENT

This work was supported by the Swiss National Science Foundation (SNSF) under project 200021 212758.

REFERENCES

[1] M. I. Sancer, R. L. McClary, and K. J. Glover, "Electromagnetic computation using parametric geometry," *Electromagnetics*, vol. 10, no. 1-2, pp. 85-103, Jan. 1990.

[2] E. Newman, P. Alexandropoulos, and E. Walton, "Polygonal plate modeling of realistic structures," *IEEE Trans. Antennas Propag.*, vol. 32, no. 7, pp. 742-747, July 1984.

[3] K. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propag.*, vol. 14, no. 3, pp. 302-307, May 1966.

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

[5] W. C. Chew, M. S. Tong, and B. Hu, *Integral Equation Methods for Electromagnetic and Elastic Waves*. San Rafael, CA: Morgan & Claypool, 2008.

[6] K. Ntokos, P. Mavrikakis, A. C. Iossifides, and T. V. Yioultsis, "A systematic approach for reconfigurable reflecting metasurface synthesis: From periodic analysis to far-field scattering," *Int. J. Electron. Commun.*, vol. 170, no. 154780, p. 154780, Oct. 2023.

[7] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*. Hoboken, NJ: John Wiley & Sons, Apr. 1998.

[8] E. M. Purcell and C. R. Pennypacker, "Scattering and absorption of light by nonspherical dielectric grains," *Astrophys. J.*, vol. 186, p. 705, Dec. 1973.

[9] B. T. Draine, "The discrete-dipole approximation and its application to interstellar graphite grains," *Astrophys. J.*, vol. 333, p. 848, Oct. 1988.

[10] O. J. F. Martin and N. B. Piller, "Electromagnetic scattering in polarizable backgrounds," *Phys. Rev. E*, vol. 58, no. 3, pp. 3909-3915, Sep. 1998.

[11] J.-M. Jin, J. L. Volakis, and J. D. Collins, "A finite-element-boundary-integral method for scattering and radiation by two- and three-dimensional structures," *IEEE Antennas Propag. Mag.*, vol. 33, no. 3, pp. 22-32, June 1991.

[12] J.-M. Jin, *The Finite Element Method in Electromagnetics*. Hoboken, NJ: John Wiley & Sons, 1993.

[13] P. P. Silvester and R. L. Ferrari, *Finite Elements for Electrical Engineers*, 3rd ed. Cambridge: Cambridge University Press, 2012.

[14] S. S. Zivanovic, K. S. Yee, and K. K. Mei, "A subgridding method for the time-domain finite-difference method to solve Maxwell's equations," *IEEE Trans. Microw. Theory Tech.*, vol. 39, no. 3, pp. 471-479, Mar. 1991.

[15] K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*. Boca Raton, FL: CRC Press, May 2018.

[16] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 1995.

[17] K. S. Yee and J. S. Chen, "The finite-difference time-domain (FDTD) and the finite-volume time-domain (FVTD) methods in solving Maxwell's equations," *IEEE Trans. Antennas Propag.*, vol. 45, no. 3, pp. 354-363, Mar. 1997.

[18] J. L. Volakis and K. Sertel, *Integral Equation Methods for Electromagnetics*. Raleigh, NC: SciTech Publishing, Jan. 2012.

[19] J. P. Kottmann and O. J. F. Martin, "Accurate solution of the volume integral equation for high-permittivity scatterers," *IEEE Trans. Antennas Propag.*, vol. 48, no. 11, pp. 1719-1726, Nov. 2000.

[20] L. Dal Negro, G. Miano, G. Rubinacci, A. Tamburrino, and S. Ventre, "A fast computation method for the analysis of an array of metallic nanoparticles," *IEEE Trans. Magn.*, vol. 45, no. 3, pp. 1618-1621, Mar. 2009.

[21] B. Gallinet, J. Butet, and O. J. F. Martin, "Numerical methods for nanophotonics: Standard problems and future challenges," *Laser Photon. Rev.*, vol. 9, no. 6, pp. 577-603, Nov. 2015.

[22] B. M. Kolundzija, "Electromagnetic modeling of composite metallic and dielectric structures," *IEEE Trans. Microw. Theory Tech.*, vol. 47, no. 7, pp. 1021-1032, July 1999.

[23] J. R. Mautz and R. F. Harrington, "H-field, E-field, and combined-field solutions for conducting bodies of revolution," *Archiv Elektronik und Uebertragungstechnik*, vol. 32, pp. 157-164, Apr. 1978.

[24] A. J. Poggio and E. K. Miller, "Chapter 4 - Integral equation solutions of three-dimensional scattering problems," in R. Mittra, editor, *Computer Techniques for Electromagnetics*, pp. 159-264, Pergamon, 1973.

[25] C. Müller, *Foundations of the Mathematical Theory of Electromagnetic Waves*. Heidelberg: Springer-Verlag, 1969.

[26] L. Medgyesi-Mitschang and J. Putnam, "Integral equation formulations for imperfectly conducting scatterers," *IEEE Trans. Antennas Propag.*, vol. 33, no. 2, pp. 206-214, Feb. 1985.

[27] A. Kishk and L. Shafai, "Different formulations for numerical solution of single or multibodies of revolution with mixed boundary conditions," *IEEE Trans. Antennas Propag.*, vol. 34, no. 5, pp. 666-673, May 1986.

[28] K. Umashankar, A. Taflove, and S. Rao, "Electromagnetic scattering by arbitrary shaped three-dimensional homogeneous lossy dielectric objects," *IEEE Trans. Antennas Propag.*, vol. 34, no. 6, pp. 758-766, June 1986.

[29] A. W. Glisson, "Electromagnetic scattering by arbitrarily shaped surfaces with impedance boundary conditions," *Radio Sci.*, vol. 27, no. 6, pp. 935-943, Nov. 1992.

[30] L. N. Medgyesi-Mitschang, J. M. Putnam, and M. B. Gedera, "Generalized method of moments for three-dimensional penetrable scatterers," *J. Opt. Soc. Am. A*, vol. 11, no. 4, p. 1383, Apr. 1994.

[31] I. Fredholm, "Sur une classe d'équations fonctionnelles," *Acta Math.*, vol. 27, pp. 365-390, 1903.

[32] S. Yan, J.-M. Jin, and Z. Nie, "Improving the accuracy of the second-kind Fredholm integral equations by using the Buffa-Christiansen functions," *IEEE Trans. Antennas Propag.*, vol. 59, no. 4, pp. 1299-1310, Apr. 2011.

[33] K. Cools, F. P. Andriulli, F. Olyslager, and E. Michielssen, "Improving the MFIE's accuracy by using a mixed discretization," in *2009 IEEE Antennas and Propagation Society International Symposium*, North Charleston, SC, June 2009.

[34] K. Cools, F. P. Andriulli, D. De Zutter, and E. Michielssen, "Accurate and conforming mixed discretization of the MFIE," *IEEE Antennas Wirel. Propag. Lett.*, vol. 10, pp. 528-531, 2011.

[35] A. M. Kern and O. J. F. Martin, "Modeling near-field properties of plasmonic nanoparticles: A surface integral approach," in *Plasmonics: Nanoimaging, Nanofabrication, and their Applications V*, San Diego, CA, Aug. 2009.

[36] C.-T. Tai, *Dyadic Green Functions in Electromagnetic Theory*. New York, NY: IEEE Press, 1994.

[37] W. C. Chew, "Integral equations," in *Waves and Fields in Inhomogenous Media*, pp. 429-509, Wiley-IEEE Press, New York, NY, 1995.

[38] B. Gallinet, A. M. Kern, and O. J. F. Martin, "Accurate and versatile modeling of electromagnetic scattering on periodic nanostructures with a surface integral approach," *J. Opt. Soc. Am. A*, vol. 27, no. 10, pp. 2261-2271, Oct. 2010.

[39] A. G. Polimeridis and T. V. Yioultsis, "On the direct evaluation of weakly singular integrals in

Galerkin mixed potential integral equation formulations," *IEEE Trans. Antennas Propag.*, vol. 56, no. 9, pp. 3011-3019, Sep. 2008.

[40] R. J. Adams, "Physical and analytical properties of a stabilized electric field integral equation," *IEEE Trans. Antennas Propag.*, vol. 52, no. 2, pp. 362-372, Feb. 2004.

[41] S. Yan, J.-M. Jin, and Z. Nie, "Accuracy improvement of the second-kind integral equations for generally shaped objects," *IEEE Trans. Antennas Propag.*, vol. 61, no. 2, pp. 788-797, Feb. 2013.

[42] P. Ylä-Oijala, S. P. Kiminki, J. Markkanen, and S. Järvenpää, "Error-controllable and well-conditioned MoM solutions in computational electromagnetics: Ultimate surface integral-equation formulation [open problems in cem]," *IEEE Antennas Propag. Mag.*, vol. 55, no. 6, pp. 310-331, Dec. 2013.

[43] R. Kress, *Linear Integral Equations*. New York, NY: Springer, Dec. 2013.

[44] D. Colton and R. Kress, "Chapter 4: Boundary-value problems for the time-harmonic Maxwell equations and the vector Helmholtz equation," in *Integral Equation Methods in Scattering Theory*, pp. 108-149, Society for Industrial and Applied Mathematics, Philadelphia, PA, Nov. 2013.

[45] X. Q. Sheng, J.-M. Jin, J. Song, W. C. Chew, and C.-C. Lu, "Solution of combined-field integral equation using multilevel fast multipole algorithm for scattering by homogeneous bodies," *IEEE Trans. Antennas Propag.*, vol. 46, no. 11, pp. 1718-1726, 1998.

[46] R. F. Harrington, "Boundary integral formulations for homogeneous material bodies," *J. Electromagn. Waves Appl.*, vol. 3, no. 1, pp. 1-15, Jan. 1989.

[47] P. Ylä-Oijala and M. Taskinen, "Well-conditioned Müller formulation for electromagnetic scattering by dielectric objects," *IEEE Trans. Antennas Propag.*, vol. 53, no. 10, pp. 3316-3323, Oct. 2005.

[48] P. Ylä-Oijala, M. Taskinen, and S. Järvenpää, "Surface integral equation formulations for solving electromagnetic scattering problems with iterative methods," *Radio Sci.*, vol. 40, no. 6, Dec. 2005.

[49] A. M. Kern and O. J. F. Martin, "Surface integral formulation for 3D simulations of plasmonic and high permittivity nanostructures," *J. Opt. Soc. Am. A*, vol. 26, no. 4, pp. 732-740, Apr. 2009.

[50] J. R. Mautz and R. F. Harrington, "Electromagnetic scattering from a homogeneous material body of revolution," *Archiv Elektronik und Uebertragungstechnik*, vol. 33, pp. 71-80, Feb. 1979.

[51] F. P. Andriulli, K. Cools, H. Bagci, F. Olyslager, A. Buffa, S. Christiansen, and E. Michielssen, "A multiplicative Calderon preconditioner for the electric field integral equation," *IEEE Trans. Antennas Propag.*, vol. 56, no. 8, pp. 2398-2412, Aug. 2008.

[52] S. Yan, J.-M. Jin, and Z. Nie, "A Comparative Study of Calderón Preconditioners for PMCHWT Equations," *IEEE Trans. Antennas Propag.*, vol. 58, no. 7, pp. 2375-2383, July 2010.

[53] B. H. Jung, T. K. Sarkar, and Y.-S. Chung, "A survey of various frequency domain integral equations for the analysis of scattering from three-dimensional dielectric objects," *Electromagn. Waves (Camb.)*, vol. 36, pp. 193-246, 2002.

[54] P. Ylä-Oijala and M. Taskinen, "Application of combined field Integral equation for electromagnetic scattering by dielectric and composite objects," *IEEE Trans. Antennas Propag.*, vol. 53, no. 3, pp. 1168-1173, Mar. 2005.

[55] R. F. Harrington, "Two-dimensional electromagnetic fields," in *Field Computation by Moment Methods*, pp. 41-61, Wiley-IEEE Press, New York, NY, 1993.

[56] G. C. Hsiao and R. E. Kleinman, "Mathematical foundations for error estimation in numerical solutions of integral equations in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 45, no. 3, pp. 316-328, Mar. 1997.

[57] J. L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*. Heidelberg: Springer-Verlag, Jan. 1972.

[58] A. Buffa, "Trace theorems on non-smooth boundaries for functional spaces related to Maxwell equations: An overview," in *Computational Electromagnetics*, pp. 23-34, Springer, Heidelberg, 2003.

[59] M. Cessenat, *Mathematical Methods in Electromagnetism: Linear Theory and Applications*. Singapore: World Scientific Publishing, Jan. 1996.

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

[61] A. Kern, "Realistic modeling of 3D plasmonic systems: A surface integral equation approach," Ph.D. thesis, Nanophotonics and Metrology Lab., EPFL, Lausanne, Switzerland, 2011.

[62] A. Buffa and S. H. Christiansen, "A dual finite element complex on the barycentric refinement," *C. R. Math. Acad. Sci. Paris*, vol. 340, no. 6, pp. 461-464, Feb. 2005.

[63] A. M. Kern and O. J. F. Martin, "Excitation and reemission of molecules near realistic plasmonic nanostructures," *Nano Lett.*, vol. 11, no. 2, pp. 482-487, Feb. 2011.

[64] L. C. Trintinalia and H. Ling, "First order triangular patch basis functions for electromagnetic scattering analysis," *J. Electromagn. Waves Appl.*, vol. 15, no. 11, pp. 1521-1537, Jan. 2001.

[65] O. Ergul and L. Gurel, "Linear-linear basis functions for MLFMA solutions of magnetic-field and combined-field integral equations," *IEEE Trans. Antennas Propag.*, vol. 55, no. 4, pp. 1103-1110, Apr. 2007.

[66] P. Ylä-Oijala, S. P. Kiminki, K. Cools, F. P. Andriulli, and S. Järvenpää, "Mixed discretization schemes for electromagnetic surface integral equations," *Int. J. Numer. Model.*, vol. 25, no. 5-6, pp. 525-540, Sep. 2012.

[67] P. Ylä-Oijala and M. Taskinen, "Calculation of CFIE impedance matrix elements with RWG and $\hat{\mathbf{n}} \times$ RWG functions," *IEEE Trans. Antennas Propag.*, vol. 51, no. 8, pp. 1837-1846, Aug. 2003.

[68] I. Hanninen, M. Taskinen, and J. Sarvas, "Singularity subtraction integral formulae for surface integral equations with RWG, rooftop and hybrid basis functions," *Prog. In Electromagn. Res.*, vol. 63, pp. 243-278, Aug. 2006.

[69] D. M. Solís, J. M. Taboada, O. Rubiños-López, and F. Obelleiro, "Improved combined tangential formulation for electromagnetic analysis of penetrable bodies," *J. Opt. Soc. Am. B*, vol. 32, no. 9, p. 1780, Sep. 2015.

[70] C. Forestiere, G. Iadarola, G. Rubinacci, A. Tamburrino, L. Dal Negro, and G. Miano, "Surface integral formulations for the design of plasmonic nanostructures," *J. Opt. Soc. Am. A*, vol. 29, no. 11, pp. 2314-2327, Nov. 2012.

[71] M. G. Araújo, J. M. Taboada, D. M. Solís, J. Rivero, L. Landesa, and F. Obelleiro, "Comparison of surface integral equation formulations for electromagnetic analysis of plasmonic nanoscatterers," *Opt. Express*, vol. 20, no. 8, pp. 9161-9171, Apr. 2012.

[72] P. Ylä-Oijala, J. Markkanen, S. Jarvenpaa, and S. P. Kiminki, "Surface and volume integral equation methods for time-harmonic solutions of Maxwell's equations (Invited Paper)," *Prog. in Electromagn. Res.*, vol. 149, pp. 15-44, Aug. 2014.

[73] B. Karaosmanoglu, A. Yilmaz, and O. Ergul, "A comparative study of surface integral equations for accurate and efficient analysis of plasmonic structures," *IEEE Trans. Antennas Propag.*, vol. 65, no. 6, pp. 3049-3057, June 2017.

[74] P. Mavrikakis and O. J. F. Martin, "Hybrid meshes for the modeling of optical nanostructures with the surface integral equation method," *J. Opt. Soc. Am. A Opt. Image Sci. Vis.*, vol. 42, no. 3, p. 315, Mar. 2025.

[75] O. Ergul and L. Gurel, "The use of curl-conforming basis functions for the magnetic-field integral equation," *IEEE Trans. Antennas Propag.*, vol. 54, no. 7, pp. 1917-1926, July 2006.

[76] E. Ubeda and J. M. Rius, "Novel monopolar MFIE MoM-discretization for the scattering analysis of small objects," *IEEE Trans. Antennas Propag.*, vol. 54, no. 1, pp. 50-57, Jan. 2006.

[77] L. Zhang, A. Deng, M. Wang, and S. Yang, "Numerical study of a novel MFIE formulation using monopolar RWG basis functions," *Comput. Phys. Commun.*, vol. 181, no. 1, pp. 52-60, Jan. 2010.

[78] X. Wan, Y. Liu, and C. Wu, "The research of magnetic-field integration method based on linear-linear basis functions," *Procedia Comput. Sci.*, vol. 183, pp. 783-790, 2021.

[79] P. Ylä-Oijala, M. Taskinen, and S. Järvenpää, "Analysis of surface integral equations in electromagnetic scattering and radiation problems," *Eng. Anal. Bound. Elem.*, vol. 32, no. 3, pp. 196-209, Mar. 2008.

[80] E. K. Miller, "A selective survey of computational electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 36, no. 9, pp. 1281-1305, 1988.

[81] W. H. Press, W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery, *Numerical Recipes in C++: The Art of Scientific Computing*, 2nd ed. Cambridge: Cambridge University Press, 2001.

[82] M. Jiang, Y. Li, L. Lei, and J. Hu, "A review on fast direct methods of surface integral equations for analysis of electromagnetic scattering from 3-D PEC objects," *Electronics (Basel)*, vol. 11, no. 22, p. 3753, Nov. 2022.

[83] G. Angiulli, M. Cacciola, S. Calcagno, D. De Carlo, C. F. Morabito, A. Sgró, and M. Versaci, "A numerical study on the performances of the flexible BiCGStab to solve the discretized E-field integral equation," *Int. J. Appl. Electromagn. Mech.*, vol. 46, no. 3, pp. 547-553, Aug. 2014.

[84] B. Carpentieri, M. Tavelli, D.-L. Sun, T.-Z. Huang, and Y.-F. Jing, "Fast iterative solution of multiple right-hand sides MoM linear systems on CPUs and GPUs computers," *IEEE Trans. Microw. Theory Tech.*, vol. 72, no. 8, pp. 4431-4444, Aug. 2024.

[85] G. Vecchi, "Loop-star decomposition of basis functions in the discretization of the EFIE," *IEEE*

Trans. Antennas Propag., vol. 47, no. 2, pp. 339-346, 1999.

[86] J.-F. Lee, R. Lee, and R. J. Burkholder, "Loop star basis functions and a robust preconditioner for EFIE scattering problems," *IEEE Trans. Antennas Propag.*, vol. 51, no. 8, pp. 1855-1863, Aug. 2003.

[87] J. F. Lee, "25 Years of Progress and Future Challenges in Fast, Iterative, and Hybrid Techniques for Large-Scale Electromagnetic Modeling," in *Proc. 25th ACES Conf.*, Monterey, CA, USA, Mar. 8–12, 2009, pp. 278–285.

[88] J.-S. Zhao and W. C. Chew, "Integral equation solution of Maxwell's equations from zero frequency to microwave frequencies," *IEEE Trans. Antennas Propag.*, vol. 48, no. 10, pp. 1635-1645, 2000.

[89] Y.-H. Chu and W. C. Chew, "Large-scale computation for electrically small structures using surface-integral equation method," *Microw. Opt. Technol. Lett.*, vol. 47, no. 6, pp. 525-530, Dec. 2005.

[90] S. H. Christiansen and J.-C. Nédélec, "A preconditioner for the electric field integral equation based on Calderon formulas," *SIAM J. Numer. Anal.*, vol. 40, no. 3, pp. 1100-1135, Jan. 2002.

[91] R. J. Adams, "Combined field integral equation formulations for electromagnetic scattering from convex geometries," *IEEE Trans. Antennas Propag.*, vol. 52, no. 5, pp. 1294-1303, May 2004.

[92] H. Contopanagos, B. Dembart, M. Epton, J. J. Ottusch, V. Rokhlin, J. L. Visher, and S. M. Wandzura, "Well-conditioned boundary integral equations for three-dimensional electromagnetic scattering," *IEEE Trans. Antennas Propag.*, vol. 50, no. 12, pp. 1824-1830, Dec. 2002.

[93] Y. Beghein, K. Cools, F. P. Andriulli, D. De Zutter, and E. Michielssen, "A Calderon multiplicative preconditioner for the PMCHWT equation for scattering by chiral objects," *IEEE Trans. Antennas Propag.*, vol. 60, no. 9, pp. 4239-4248, Sep. 2012.

[94] J. M. Song, C. C. Lu, W. C. Chew, and S. W. Lee, "Fast Illinois solver code (FISC)," *IEEE Antennas Propag. Mag.*, vol. 40, no. 3, pp. 27-34, June 1998.

[95] L. Greengard and V. Rokhlin, "A fast algorithm for particle simulations," *J. Comput. Phys.*, vol. 73, no. 2, pp. 325-348, Dec. 1987.

[96] V. Rokhlin, "Rapid solution of integral equations of scattering theory in two dimensions," *J. Comput. Phys.*, vol. 86, no. 2, pp. 414-439, Feb. 1990.

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

[98] O. Ergul and L. Gurel, *The Multilevel Fast Multipole Algorithm (MLFMA) for Solving Large-Scale Computational Electromagnetics Problems*. Hoboken, NJ: John Wiley & Sons, June 2014.

[99] J. Guan, S. Yan, and J.-M. Jin, "An OpenMP-CUDA implementation of multilevel fast multipole algorithm for electromagnetic simulation on multi-GPU computing systems," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3607-3616, July 2013.

[100] M.-L. Yang, B.-Y. Wu, H.-W. Gao, and X.-Q. Sheng, "A ternary parallelization approach of MLFMA for solving electromagnetic scattering problems with over 10 billion unknowns," *IEEE Trans. Antennas Propag.*, vol. 67, no. 11, pp. 6965-6978, Nov. 2019.

[101] W.-J. He, X.-W. Huang, M.-L. Yang, and X.-Q. Sheng, "Massively parallel multilevel fast multipole algorithm for extremely large-scale electromagnetic simulations: A review," *Electromagn. Waves (Camb.)*, vol. 173, pp. 37-52, 2022.

[102] C.-F. Wang, F. Ling, J. Song, and J.-M. Jin, "Adaptive integral solution of combined field integral equation," *Microw. Opt. Technol. Lett.*, vol. 19, no. 5, pp. 321-328, Dec. 1998.

[103] O. P. Bruno and L. A. Kunyansky, "A fast, high-order algorithm for the solution of surface scattering problems: Basic implementation, tests, and applications," *J. Comput. Phys.*, vol. 169, no. 1, pp. 80-110, May 2001.

[104] S.-Q. Li, Y. Yu, C. H. Chan, K. F. Chan, and L. Tsang, "A sparse-matrix/canonical grid method for analyzing densely packed interconnects," *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 7, pp. 1221-1228, July 2001.

[105] S. Gedney, A. Zhu, W.-H. Tang, G. Liu, and P. Petre, "A fast, high-order quadrature sampled pre-corrected fast-Fourier transform for electromagnetic scattering," *Microw. Opt. Technol. Lett.*, vol. 36, no. 5, pp. 343-349, Mar. 2003.

[106] Fasenfest, Capolino, Wilton, Jackson, and Champagne, "A fast MoM solution for large arrays: Green's function interpolation with FFT," *IEEE Antennas Wirel. Propag. Lett.*, vol. 3, pp. 161-164, 2004.

[107] A. Heldring, J. M. Tamayo, C. Simon, E. Ubeda, and J. M. Rius, "Sparsified adaptive cross approximation algorithm for accelerated method of moments computations," *IEEE Trans. Antennas Propag.*, vol. 61, no. 1, pp. 240-246, Jan. 2013.

[108] P. G. Martinsson and V. Rokhlin, "A fast direct solver for scattering problems involving elongated

structures," *J. Comput. Phys.*, vol. 221, no. 1, pp. 288-302, Jan. 2007.

[109] J. Shaeffer, "Direct solve of electrically large integral equations for problem sizes to 1 M unknowns," *IEEE Trans. Antennas Propag.*, vol. 56, no. 8, pp. 2306-2313, Aug. 2008.

[110] E. Candès, L. Demanet, and L. Ying, "A fast butterfly algorithm for the computation of Fourier integral operators," *Multiscale Model. Simul.*, vol. 7, no. 4, pp. 1727-1750, Jan. 2009.

[111] M. A. E. Bautista, M. A. Francavilla, P.-G. Martinsson, and F. Vipiana, "$\mathcal{O}(N)$ nested skeletonization scheme for the analysis of multiscale structures using the method of moments," *IEEE J. Multiscale Multiphys. Comput. Tech.*, vol. 1, no. 1, pp. 139-150, 2016.

[112] S. Kapur and D. Long, "IES^3: A fast integral equation solver for efficient 3-dimensional extraction," in *Proceedings of IEEE International Conference on Computer Aided Design (ICCAD)*, San Jose, CA, USA, pp. 448-455, Nov. 9-13, 1997.

[113] E. Michielssen and A. Boag, "A multilevel matrix decomposition algorithm for analyzing scattering from large structures," *IEEE Trans. Antennas Propag.*, vol. 44, no. 8, pp. 1086-1093, 1996.

[114] A. Heldring, J. M. Rius, J. M. Tamayo, J. Parrón, and E. Ubeda, "Multiscale compressed block decomposition for fast direct solution of method of moments linear system," *IEEE Trans. Antennas Propag.*, vol. 59, no. 2, pp. 526-536, Feb. 2011.

[115] Y. Brick, V. Lomakin, and A. Boag, "Fast direct solver for essentially convex scatterers using multilevel non-uniform grids," *IEEE Trans. Antennas Propag.*, vol. 62, no. 8, pp. 4314-4324, Aug. 2014.

[116] W. Hackbusch, "A sparse matrix arithmetic based on $\mathscr{H}$ -matrices. Part I: Introduction to $\mathscr{H}$ -matrices," *Computing*, vol. 62, no. 2, pp. 89-108, Apr. 1999.

[117] W. Hackbusch and B. N. Khoromskij, "A sparse $\mathscr{H}$-matrix arithmetic," *Computing*, vol. 64, no. 1, pp. 21-47, Feb. 2000.

[118] S. Börm, "$\mathscr{H}^2$-matrices - Multilevel methods for the approximation of integral operators," *Comput. Vis. Sci.*, vol. 7, no. 3-4, pp. 173-181, Oct. 2004.

[119] W. Chai and D. Jiao, "Dense matrix inversion of linear complexity for integral-equation-based large-scale 3-D capacitance extraction," *IEEE Trans. Microw. Theory Tech.*, vol. 59, no. 10, pp. 2404-2421, Oct. 2011.

[120] J.-G. Wei, Z. Peng, and J.-F. Lee, "A fast direct matrix solver for surface integral equation methods for electromagnetic wave scattering from nonpenetrable targets," *Radio Sci.*, vol. 47, no. 5, Oct. 2012.

[121] M. Ma and D. Jiao, "Accuracy directly controlled fast direct solution of general $\mathscr{H}^2$ -matrices and its application to solving electrodynamic volume integral equations," *IEEE Trans. Microw. Theory Tech.*, vol. 66, no. 1, pp. 35-48, Jan. 2018.

[122] H. Guo, J. Hu, H. Shao, and Z. Nie, "Hierarchical matrices method and its application in electromagnetic integral equations," *Int. J. Antennas Propag.*, vol. 2012, pp. 1-9, 2012.

[123] H. Guo, Y. Liu, J. Hu, and E. Michielssen, "A butterfly-based direct integral-equation solver using hierarchical LU factorization for analyzing scattering from electrically large conducting objects," *IEEE Trans. Antennas Propag.*, vol. 65, no. 9, pp. 4742-4750, Sep. 2017.

[124] S. Wang, "Quasi-block-Cholesky factorization with dynamic matrix compression for fast integral-equation simulations of large-scale human body models," *Proc. IEEE Inst. Electr. Electron. Eng.*, vol. 101, no. 2, pp. 389-400, Feb. 2013.

[125] H. Guo, Y. Liu, J. Hu, and E. Michielssen, "A butterfly-based direct solver using hierarchical LU factorization for Poggio-Miller-Chang-Harrington-Wu-Tsai equations," *Microw. Opt. Technol. Lett.*, vol. 60, no. 6, pp. 1381-1387, June 2018.

[126] B. M. Notaros, B. D. Popovic, J. P. Weem, R. A. Brown, and Z. Popovic, "Efficient large-domain MoM solutions to electrically large practical EM problems," *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 1, pp. 151-159, 2001.

[127] M. Djordjevic and B. M. Notaros, "Double higher order method of moments for surface integral equation modeling of metallic and dielectric antennas and scatterers," *IEEE Trans. Antennas Propag.*, vol. 52, no. 8, pp. 2118-2129, Aug. 2004.

[128] B. M. Notaros, "Higher order frequency-domain computational electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 56, no. 8, pp. 2251-2276, Aug. 2008.

[129] B. Michiels, J. Fostier, I. Bogaert, and D. De Zutter, "Full-wave simulations of electromagnetic scattering problems with billions of unknowns," *IEEE Trans. Antennas Propag.*, vol. 63, no. 2, pp. 796-799, Feb. 2015.

[130] N. V. Nair and B. Shanker, "Generalized method of moments: A novel discretization technique for integral equations," *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 2280-2293, June 2011.

[131] Z. Peng, K.-H. Lim, and J.-F. Lee, "A discontinuous Galerkin surface integral equation method for electromagnetic wave scattering from

nonpenetrable targets," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3617-3628, July 2013.

[132] P. Ylä-Oijala, S. P. Kiminki, K. Cools, F. P. Andriulli, and S. Järvenpää, "Stable discretization of combined source integral equation for scattering by dielectric objects," *IEEE Trans. Antennas Propag.*, vol. 60, no. 5, pp. 2575-2578, May 2012.

Parmenion S. Mavrikakis was born in Thessaloniki, Macedonia, Greece, in 1998. He received the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki (AUTh), Thessaloniki, Macedonia, Greece, in 2022.

In 2023, he joined the Nanophotonics and Metrology Laboratory (NAM), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, as a Ph.D. student. He is currently conducting research on the development of numerical techniques for the solution of electromagnetic scattering problems on plasmonic, dielectric and hybrid systems. His research interests are in the areas of computational electromagnetics, metasurfaces, microwaves, photonics and plasmonics.

Olivier J. F. Martin is currently a Professor of nanophotonics and optical signal processing with the Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland, where he is also the Head of the Nanophotonics and Metrology Laboratory (NAM). He conducts a comprehensive research that combines the development of numerical techniques for the solution of Maxwell's equations with advanced nanofabrication and experiments on plasmonic systems. He has authored over 300 journal articles and holds a number of patents and invention disclosures. Applications of his research include metasurfaces, nonlinear optics, biosensing, heterogeneous catalysis, security features, and optical forces at the nanoscale. In 2016, he received an ERC Advanced Grant on the utilization of plasmonic forces to fabricate nanostructures.

Superior Accuracy of the Normally-integrated MFIE Compared to the Traditional MFIE

Andrew F. Peterson and Malcolm M. Bibby

School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332-0250, USA
peterson@ece.gatech.edu, mbibby@ece.gatech.edu

Abstract – An alternative method of moments discretization of the magnetic field integral equation (MFIE) uses testing functions inside the target and in a plane normal to the target surface. This approach is adapted to targets modeled with flat-faceted patches. A comparison with traditional numerical solutions of the MFIE that use testing functions on the target surface shows that the normally-integrated MFIE formulation produce far fields that are more accurate than those obtained from the traditional MFIE. The alternate approach can be made free from internal resonances and that approach is often more accurate than the combined field integral equation.

Index Terms – Electromagnetic scattering, method of moments, numerical techniques, radar cross section, scattering cross section.

I. INTRODUCTION

Integral equations such as the electric-field equation (EFIE) and magnetic field integral equation (MFIE) are the foundation for many numerical solution techniques in electromagnetics, especially open-region problems such as radiation and scattering applications involving perfectly conducting materials. The most common discretization procedure used with the EFIE involves Rao-Wilton-Glisson (RWG) basis and testing functions [1], which converts the "strong" form of the EFIE into a "weak" equation where the degree of the operator derivative has been reduced by the testing function. A similar discretization of the MFIE does not produce a "weak" equation since one of the MFIE derivatives is normal to the surface being discretized and not affected by the tangential testing function. The MFIE has other restrictions: it is not applicable to open targets and may fail under certain conditions for multiply connected targets [2]. Both equations many fail at internal resonance frequencies [3]. In addition, the MFIE is thought to be more sensitive to discontinuities, such as those introduced by flat-faceted models of curved surfaces [4].

The accuracy of the far fields (and scattering cross section) produced by traditional MFIE discretizations is substantially worse than that produced by the EFIE [5–6]. When the EFIE and MFIE are combined together to form the combined field integral equation (CFIE) to avoid internal resonance failures, the accuracy of the CFIE is degraded by the underlying MFIE accuracy and seldom equals that of the EFIE away from resonances.

In [7–8], an alternative discretization was introduced for the MFIE, involving testing functions that are inside the target and in a plane normal to the surface. These functions can absorb both derivatives arising from the curl operator. The approach, labeled the normally-integrated MFIE or NIMFIE, produces a true "weak" equation different from that of the traditional tangentially-tested MFIE. In [7–9], the NIMFIE was demonstrated for several smooth targets, using high order basis functions and perfect models of the curved target surfaces. In order to employ the NIMFIE for more complex targets modeled with flat patch models, a new discretization is proposed with RWG basis functions representing the current and testing functions with support that may be divided into two non-overlapping domains inside the target. Results from the new approach will be compared with those produced by a conventional MFIE discretization (also using RWG basis functions but with RWG testing functions located on the target surface). In addition, [8] demonstrated that, by including an exponential phase in the testing functions, the original NIMFIE equations could be made free of internal resonances. Consequently, we use the "resonant free" approach in the following and also compare the results to those of the CFIE. Preliminary results from this study were presented in [10–11].

II. NIMFIE FORMULATION

The NIMFIE formulation is based on the condition that the total magnetic field vanishes inside a perfectly conducting target. Therefore, the incident and scattered magnetic fields satisfy:

$$\bar{T} \bullet \bar{H}^{inc} + \bar{T} \bullet \bar{H}^{s} = 0, \tag{1}$$

where the incident field is that of the primary source in the absence of the target, and the scattered field is the field of the equivalent current density $\bar{J}$ on the target surface, also computed in the absence of the target. The test-

Submitted On: July 23, 2024
Accepted On: March 31, 2025

https://doi.org/10.13052/2025.ACES.J.400403

ing function $\bar{T}$ occupies a plane normal to the surface of the target and is a vector perpendicular to that plane. The scattered field is obtained from:

$$\bar{H}^s = \nabla \times \bar{A}, \tag{2}$$

where $\bar{A}$ is the magnetic vector potential function:

$$\bar{A}(\bar{r}) = \iint \bar{J}(s',t') \frac{e^{-jkR}}{4\pi R}\, ds'dt'. \tag{3}$$

The integral in (3) is over the target surface, $k = \omega\sqrt{\mu_0\varepsilon_0}$ is the wavenumber of the background medium, ω is the radian frequency, and μ_0 and ε_0 are the permeability and permittivity of the background medium. The parameter R is the distance from the point of integration to the point where the field is computed.

For targets represented by flat-faceted triangular patches, the basis functions representing the current density are chosen to be RWG functions that straddle pairs of patches and interpolate to the current density flowing across edges of the model [1]. For convenience each testing function will also be associated with an edge. The test function domain associated with two patches straddling an edge p is inside the target, beginning below the centroid of one patch (A) and terminating beneath the centroid of an adjacent patch (B), as depicted in Fig. 1. For patch pairs that are bent to represent curved or non-planar parts of surfaces, the testing domain in the present work is separated into two non-overlapping rectangular regions as depicted in the side view shown in Fig. 1 (b). Let (s, t, n) denote a local right-handed coordinate system associated with a patch, with variables s and t tangential to the patch and n in the outward normal direction. In the "internal resonance free" NIMFIE formulation, the part of the test function below one of the two patches is:

$$\bar{T} = \hat{s}\, p(t;t1,t2)\, e^{-j\varphi(n2-n)} p(n;n1,n2), \tag{4}$$

where $p(t;t1,t2)$ denotes a windowing function:

$$p(t;t1,t2) = \begin{cases} 1 & t1 < t < t2 \\ 0 & otherwise \end{cases}, \tag{5}$$

and the rectangular domain is the region $t1<t<t2$, $n1<n<n2$. The phase factor:

$$\varphi = \frac{\pi}{2}\frac{1}{n2-n1}, \tag{6}$$

provides a 90-degree phase progression across the test function domain, to suppress internal resonances in a manner similar to that of the dual surface integral equations [12].

The NIMFIE equation associated with the complete contribution from the basis function associated with edge q and the two parts of the test function associated with edge p can be obtained by inserting (4) and its complement into the relation:

$$-\iint \bar{T} \bullet \bar{H}^{inc} = \iint \bar{T} \bullet \nabla \times \bar{A}, \tag{7}$$

where the integrations are over the test domain inside the target. By carrying out any integrations that cancel

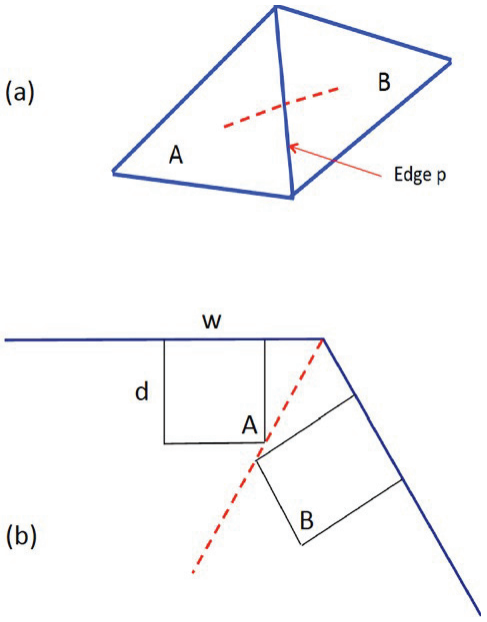


Fig. 1. (a) Top and (b) side views of a patch pair showing the two parts of the domain of the testing function when the patches are non-planar.

derivatives arising from the curl operation, we obtain the resulting equation:

$$\begin{aligned}
&-\int_{n1A}^{n2A}\int_{t1A}^{t2A} e^{-j\varphi(n2A-n)}\, \hat{s}_{pA} \bullet \bar{H}^{inc} \\
&-\int_{n1B}^{n2B}\int_{t1B}^{t2B} e^{-j\varphi(n2B-n)}\, \hat{s}_{pB} \bullet \bar{H}^{inc} \\
&= \int_{n1A}^{n2A} e^{-j\varphi(n2A-n)} \hat{n}_{pA} \bullet \bar{A}\Big|_{t1A}^{t2A} \\
&+ \int_{n1B}^{n2B} e^{-j\varphi(n2B-n)} \hat{n}_{pB} \bullet \bar{A}\Big|_{t1B}^{t2B} \\
&- \int_{t1A}^{t2A} \left(e^{-j\varphi(n2A-n)} \hat{t}_{pA} \bullet \bar{A}\right)\Big|_{n1A}^{n2A} \\
&- \int_{t1B}^{t2B} \left(e^{-j\varphi(n2B-n)} \hat{t}_{pB} \bullet \bar{A}\right)\Big|_{n1B}^{n2B} \\
&+ j\varphi_A \int_{n1A}^{n2A}\int_{t1A}^{t2A} e^{-j\varphi(n2A-n)}\, \hat{t}_{pA} \bullet \bar{A} \\
&+ j\varphi_B \int_{n1B}^{n2B}\int_{t1B}^{t2B} e^{-j\varphi(n2B-n)}\, \hat{t}_{pB} \bullet \bar{A}
\end{aligned}, \tag{8}$$

where the integration limits such as "$t1A$" are those associated with the part of the test domain under patch "A" and limits such as "$t1B$" are those associated with the test domain under patch "B" (Fig. 1). Similarly, unit vectors exhibit an "A" or "B" subscript to denote which test patch (A or B) they are associated with. Equation (8) is slightly different from those in [8] due to the vector nature of the RWG basis and the two distinct domains associated with the testing function.

Observe that there are no derivatives applied to the magnetic vector potential function in (8), confirming that this orientation of the test function absorbs all the derivatives arising from the curl operation and produces a truly "weak" form of the integral equation. For numerical implementation, the integrals in (8) are no more difficult to compute than those arising from the traditional EFIE discretizations.

Ideally, to suppress internal resonances, the test function domain should extend to a quarter wavelength depth d within the target [12]. For thinner parts of targets where the thickness is less than a half wavelength beneath a particular patch, d is reduced to half the available depth. In addition, the length w of the test domain along the variable t under each patch is initially set to the distance from the centroid to the edge but is reduced if the patch pair is bent as illustrated in Fig. 1 (b). In the present work, for interior angles greater than 135 degrees, w and d are reduced by a common factor to avoid overlap with the other part of the test domain, as depicted in Fig. 1 (b). For interior angles smaller than 135 degrees (targets with sharper bends), the constraint w=d is imposed and both dimensions are reduced to avoid overlap. In addition, to avoid large disparities in test domain size from cell to cell, adjustments in surrounding cells are made to ensure that test domain dimensions do not differ by more than a factor of 2 for test domains associated with the same patch or adjacent patches. The preprocessing associated with these computations is relatively small.

III. NUMERICAL RESULTS

As an initial example, Fig. 2 shows the error in the bistatic scattering cross section (SCS) for a perfectly conducting sphere with ka=2π, where a is the sphere radius. Flat-faceted models of spheres with the correct surface area were used. The 2-norm error in the scattering cross section is defined:

$$E = \frac{\sqrt{\frac{1}{N_{angles}} \sum_{m=1}^{N_\theta} \sum_{n=1}^{N_\phi} \sin\theta_m \, |\sigma_{exact}(\theta_m, \phi_n) - \sigma_{numerical}(\theta_m, \phi_n)|^2}}{\sigma_{exact,forward}} \tag{9}$$

The SCS error was averaged over a 5-degree grid in spherical angles (θ, ϕ), and therefore N_θ and N_ϕ assume values 37 and 73, respectively, and N_{angles}=2701. Results are presented for the EFIE, MFIE, CFIE, and NIMFIE approaches, where all used RWG basis functions and all but NMIFIE used RWG testing functions on the target surface. The CFIE used equal weighting between the EFIE and MFIE parts. These results show that, for densities of 50-200 unknowns/λ^2, where λ is the wavelength, the EFIE and the NIMFIE produce SCS results that are more than an order of magnitude more accurate than those produced by the traditional MFIE or CFIE. (In these and most other results, the EFIE is expected to exhibit more accurate far fields than the NIMFIE because of the variational superconvergence associated with its method of moments discretization [13].)

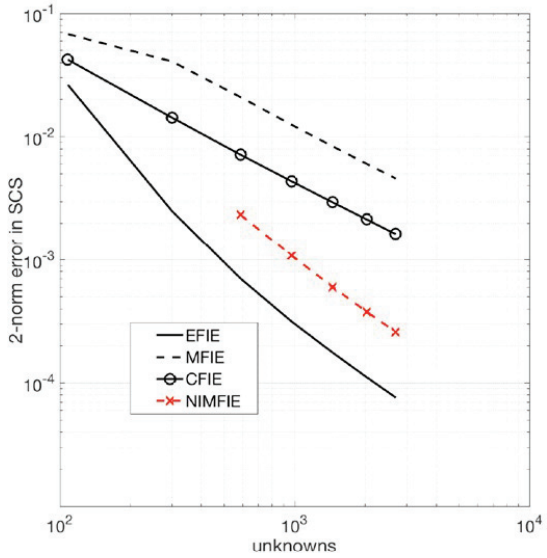


Fig. 2. Error in the SCS for a sphere with ka=2π.

We next consider several examples whose surfaces contain sharp or abrupt bends and tips. Figure 3 shows the magnitude of the surface current density for a perfectly conducting missile target [14] with surface area $25\lambda^2$, obtained from the EFIE, MFIE, and NIMFIE equations with a 2592-cell model. The target is illuminated nose-on with a horizontally-polarized incident electric field. This target has two fin-like protrusions containing 90-degree bends. For this example, the NIMFIE result exhibits qualitative agreement with the EFIE result, while the MFIE current shows minor differences.

Several measures of the error are reported in Table 1, including the error in the currents obtained by a comparison with a higher order solution of the EFIE, the tangential electric-field residual error [15], and the SCS error compared to a higher order EFIE reference result. The current density error is computed using:

$$J_{err} = \frac{\sqrt{\frac{1}{A_{total}} \sum_{n=1}^{N_{cells}} \left| \bar{J}_{n,\,ref} - \bar{J}_{n,\,RWG} \right|^2 A_n}}{2\,|H^{inc}_{\max}|}, \tag{10}$$

where A_n is the area of cell n. These results suggest that all four approaches produce somewhat similar accuracy

Table 1: 2-norm errors in the results for the missile

	J Error	E-residual	SCS Error
MFIE	17%	27%	3.3%
CFIE	18	19	2.5
NIMFIE	17	22	2.0
EFIE	19	17	0.15

in the current density, with a 2-norm error around 20%, and that the NIMFIE is slightly better than the MFIE in SCS accuracy.

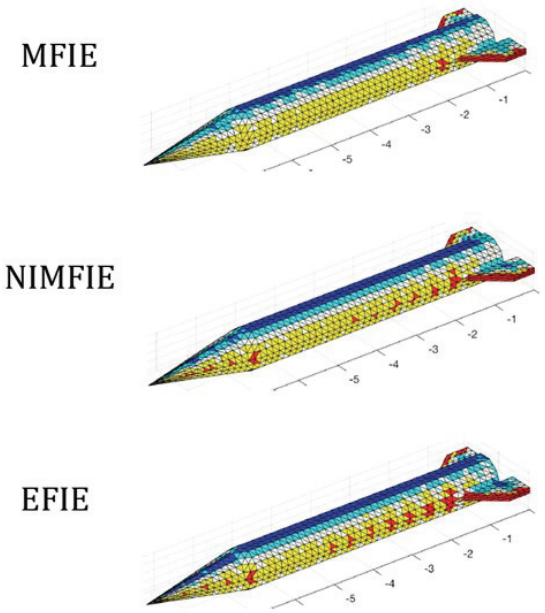


Fig. 3. Current density magnitude on a perfectly conducting missile target. Red denotes the largest magnitudes, followed by yellow, white, light blue, and dark blue, on a non-logarithmic scale. The target is illuminated nose-on with a horizontally-polarized incident electric field. The scales are in wavelength.

Figure 4 shows the surface current magnitude on a perfectly conducting cone-sphere with a total length of 5.66λ, a cone length-to-radius ratio of 4:1, and a total surface area of $25\lambda^2$, for a vertically-polarized electric field incident upon the sharp tip end of the target. A 2812-cell model is used. The NIMFIE result exhibits agreement with the EFIE result, while the MFIE result shows some differences. Various error measures are reported in Table 2. The overall errors are smaller for the cone-sphere than the missile, and the SCS error obtained from the NIMFIE results are significantly smaller than those of the MFIE and CFIE.

Table 2: 2-norm errors for the cone-sphere target

	J Error	E-residual	SCS Error
MFIE	12%	17%	1.5%
CFIE	12	14	0.80
NIMFIE	11	14	0.24
EFIE	13	13	0.030

Figure 5 shows the magnitude of the surface current on a perfectly conducting "Arrow" target [16] with surface area of $25\lambda^2$ and 2002 cells. The bottom of the target is flat, with interior angles of only 36 and 53 degrees around the front and side edges, respectively. The NIMFIE and EFIE results exhibit reasonable visual agreement, while the MFIE result shows differences. Table 3 reports various 2-norm errors. For this target, the NIMFIE SCS error is smaller than that of the MFIE or CFIE, and almost as low as the EFIE SCS error.

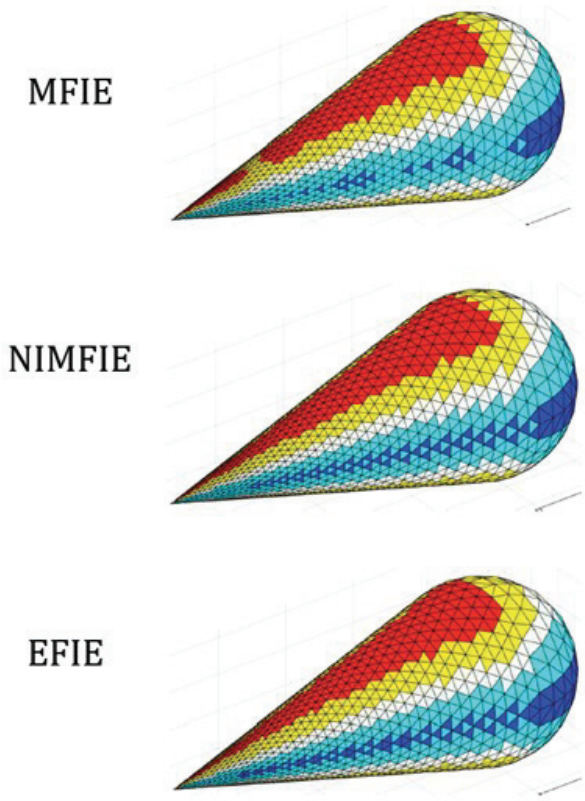


Fig. 4. Current density magnitude induced on a perfectly conducting cone-sphere target of surface area $25\lambda^2$, by a wave with a vertically polarized electric field incident upon the tip end of the target. Red denotes the largest magnitudes, followed by yellow, white, light blue, and dark blue, on a non-logarithmic scale.

Table 3: 2-norm errors for the Arrow target

	J Error	E-residual	SCS Error
MFIE	21%	36%	4.1%
CFIE	19	20	2.1
NIMFIE	17	25	0.78
EFIE	19	21	0.67

Figure 6 shows the magnitude of the current density on a cube target with cube edge length 1.58λ. This target happens to be internally resonant for the MFIE,

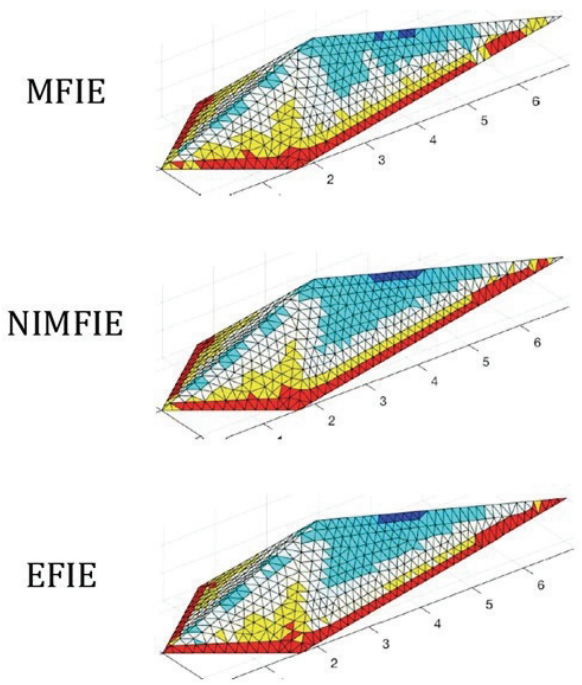


Fig. 5. Current density magnitude induced on a perfectly conducting "Arrow" target [16] with surface area of $25\lambda^2$. The target is illuminated nose-on with a horizontally-polarized incident electric field. The scales are in wavelength. Red denotes the largest magnitudes, followed by yellow, white, light blue, and dark blue, on a non-logarithmic scale.

Table 4: 2-norm errors for the 2700-cell cube target

	J Error	E-residual	SCS Error
MFIE	42%	83%	8.4%
CFIE	9	9	0.19
NIMFIE	9	11	0.19
EFIE	15	8	0.039

and the MFIE result is clearly incorrect. The NIMFIE result for the current magnitude obtained with a 2700-cell model exhibits reasonable agreement with a CFIE result obtained with a 4800-cell model. Table 4 shows several measures of the 2-norm error, obtained from 2700-cell models for the four approaches. For this example, the "resonant-free" NIMFIE and CFIE produce similar error levels in both current and SCS.

Table 5 summarizes the results of the preceding examples and several others in a different manner, by reporting the ratio of the MFIE SCS 2-norm errors to those of the other formulations. The additional targets include a second cube illuminated face-on, the Arrow illuminated by a vertically-polarized incident electric field, and two almond targets illuminated by waves incident on their blunt ends. These results show the disparity between the EFIE and MFIE SCS accuracy in practice. They also suggest that the NIMFIE formulation consistently outperforms the MFIE for SCS error, and almost always produces more accurate SCS results than the CFIE.

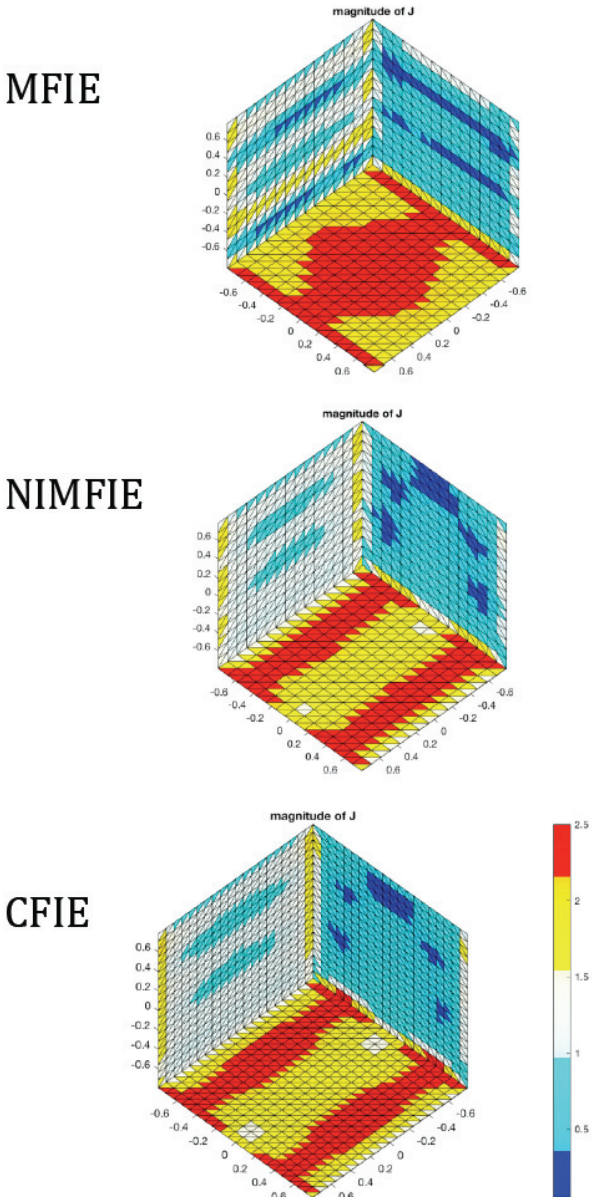


Fig. 6. Current density magnitude induced on a perfectly conducting cube target by a wave normally incident on one of the faces. The MFIE and NIMFIE results are obtained with 2700-cell models, while the CFIE result is obtained using a model with 4800 cells. The color scale is non-logarithmic and is a fraction of twice the incident magnetic field.

Table 5: Ratio of MFIE SCS error to other equation's SCS error

Target	Area (λ^2)	Density (Cells/λ^2)	CFIE	NIMFIE	EFIE
sphere	12.6	108	2.8	16	55
cube #1	15.015	180	44	44	212
almond #1, H-pol	21.8	118	2.5	2.9	28
cube #2	24	113	7.2	7.9	23
missile	25	104	1.3	1.7	22
4:1 cone-sphere	25	112	1.9	6.3	52
Arrow, H-pol	25	80	1.9	5.3	6.2
Arrow, V-pol	25	80	1.5	3.3	5.4
almond #2, V-pol	44.5	104	1.1	9.8	21

Table 6 shows the matrix condition numbers reported by the matrix solver for the various approaches and the preceding examples. The CFIE has the lowest (best) condition numbers, while the NIMFIE condition numbers are the same order as those of the EFIE. The data reflect the fact that the "cube #1" target is internally resonant for the MFIE (and apparently close to an internal resonance for the EFIE) and also suggests that there are irregularities associated with the model used for the cone-sphere, probably near the tip.

Table 6: Condition numbers for the cases in Table 5

Target	MFIE	CFIE	NIMFIE	EFIE
sphere	57	21	235	125
cube #1	1918	26	473	9293
almond #1	276	23	469	417
cube #2	91	26	463	143
missile	971	106	1509	1650
cone-sphere	7880	817	13320	12749
Arrow	592	54	2240	1905
almond #2	550	37	846	2170

IV. CONCLUSION

The NIMFIE approach has been implemented with flat-patch models and RWG basis functions, and results for a variety of targets with bends in their surfaces were compared to the traditional MFIE and the CFIE. For these targets, the NIMFIE SCS accuracy is consistently better than that produced by the MFIE and is usually better than the accuracy of the CFIE. These results support the hypothesis that the traditional MFIE discretization, using the "strong" MFIE operator, is more sensitive to surface discontinuities than the "weak" NIMFIE approach. As implemented here, the NIMFIE is free from internal resonances, and it appears to offer advantages over the traditional MFIE formulation.

REFERENCES

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

[2] A. E. Ofluoglu, T. Ciftci, and O. Ergul, "Magnetic field integral equation," *IEEE Antennas and Propagation Magazine*, vol. 57, no. 4, pp. 134-142, 2015.

[3] A. F. Peterson, "The interior resonance problem associated with surface integral equations of electromagnetics: Numerical consequences and a survey of remedies," *Electromagnetics*, vol. 10, no. 3, pp. 293-312, 1990.

[4] A. F. Peterson and M. M. Bibby, *An Introduction to the Locally-corrected Nyström Method*. San Raphael: Morgan and Claypool Synthesis Lectures, 2010.

[5] O. Ergul and L. Gurel, "Improving the accuracy of the magnetic field integral equation with linear-linear basis functions," *Radio Science*, vol. 41, no. 4, RS4004, 2006.

[6] A. F. Peterson, "Observed baseline convergence rates and superconvergence in the scattering cross section obtained from numerical solutions of the MFIE," *IEEE Trans. Antennas Propagat*., vol. 56, no. 11, pp. 3510-3515, Nov. 2008.

[7] M. M. Bibby and A. F. Peterson, "Elimination of the derivatives from the conventional MFIE operator," in *Proceedings of the 22nd Annual Review of Progress in Applied Computational Electromagnetics*, Miami, FL, pp. 359-364, Mar. 2006.

[8] M. M. Bibby, C. M. Coldwell, and A. F. Peterson, "Normally-integrated magnetic field integral equations for electromagnetic scattering," *IEEE Trans. Antennas Propagat*., vol. 55, no. 9, pp. 2530-2536, 2007.

[9] M. M. Bibby, C. M. Coldwell, and A. F. Peterson, "A high order numerical investigation of electromagnetic scattering from a torus and a circular loop," *IEEE Trans. Antennas Propagat*., vol. 61, no. 7, pp. 3656-3661, 2013.

[10] A. F. Peterson and M. M. Bibby, "Performance of the normally-integrated magnetic field integral equation for flat faceted surfaces," in *International*

Conference on Electromagnetics in Advanced Applications (ICEAA 23), Venice, Italy, Oct. 2023.

[11] A. F. Peterson and M. M. Bibby, "Superior far field accuracy of the normally-integrated MFIE compared to the conventional MFIE for flat-faceted targets," in *International Applied Computational Electromagnetics Society (ACES) Symposium*, Orlando, FL, May 2024.

[12] M. B. Woodworth and A. D. Yaghjian, "Multiwavelength three-dimensional scattering with dual-surface integral equations," *J. Opt. Soc. Am. A.*, vol. 11, no. 4, pp. 1399-1413, Apr. 1994.

[13] A. F. Peterson, D. R. Wilton, and R. E. Jorgenson, "Variational nature of Galerkin and non-Galerkin moment method solutions," *IEEE Trans. Antennas Propagat.*, vol. 44, no. 4, pp. 500-503, Apr. 1996.

[14] S. K. Kim and A. F. Peterson, "Adaptive *h*-refinement for the RWG based EFIE," *IEEE J. Multiscale and Multiphysics Comp. Tech.*, vol. 3, pp. 58-65, June 2018.

[15] A. F. Peterson, "Integral equation residuals for error estimation and internal resonance detection," *IEEE Trans. Antennas Propagat.*, vol. 71, no. 12, pp. 9326-9333, Dec. 2023.

[16] M. E. Kowalski, B. Singh, L. C. Kempel, K. D. Trott, and J.-M. Jin, "Application of the integral equation-asymptotic phase (IE-AP) method to three-dimensional scattering," *J. Electromagnetic Waves Appl.*, vol. 15, pp. 885-900, July 2001.

Andrew F. Peterson received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from the University of Illinois, Urbana-Champaign, USA. Since 1989, he has been a member of the faculty of the School of Electrical and Computer Engineering at the Georgia Institute of Technology, where he is a full Professor. Within ACES, he has served at various times as a member of the Board of Directors, the Finance Committee Chair, the Publications Committee Chair, and the President. He was elevated to ACES Fellow in 2008.

Malcolm M. Bibby received the B.Eng. and Ph.D. degrees in Electrical Engineering from the University of Liverpool, UK, in 1962 and 1965, respectively, and an MBA from the University of Chicago, USA. He is currently an Adjunct Professor in ECE at Georgia Tech. He has been interested in the numerical aspects of antenna design and electromagnetics for more than 40 years.

Advanced Physical Optics-inspired Support Vector Regression for Efficient Modeling of Target RCS

Chenge Shi[1], Rui Cai[1], Wei Dong[1], and Donghai Xiao[2]

[1]AVIC Xi'an Aircraft Industry Group Company Ltd.
Xi'an, Shaanxi 710089, China
Chen_geShi7@126.com, cairui1201@126.com, 983683251@qq.com

[2]Xidian University
Xi'an, Shaanxi 710071, China
xiaodonghai@xidian.edu.cn

***Abstract* –** This paper proposes an advanced physical optics-inspired support vector regression (APOI-SVR) for efficiently modeling the radar cross section (RCS) of conducting targets. Specifically, an improved physical optics-inspired kernel function is newly proposed by introducing two angular frequency parameters, thereby enhancing the capability of characterizing the various fluctuation patterns in RCS with respect to observation angles. Furthermore, considering the critical role of data preprocessing in facilitating the model's ability to learn the underlying RCS patterns accurately, a physics-based data preprocessing method is introduced. Numerical validations based on two exemplary targets demonstrate that APOI-SVR effectively reduces the predictive root mean square error (RMSE) by over 24.7% compared with the benchmark model. Afterward, APOI-SVR is adopted to quickly establish the RCS feature map of an aircraft model, the results show that it is comparable to numerical simulations in accuracy but less than one-tenth in time cost, indicating the practicality of APOI-SVR for efficiently analyzing the RCS characteristics of targets.

***Index Terms* –** Angular frequency parameter, data preprocessing, physical optics, radar cross section, support vector regression.

I. INTRODUCTION

In the domain of electromagnetic scattering research, the accurate prediction of the radar cross section (RCS) of a target is essential for target recognition and tracking [1, 2], and also exerts a profound influence on a variety of modern applications across aerospace and civilian domains [3–5]. Computational electromagnetic (CEM) methods, such as the finite element method (FEM) [6, 7] and the finite difference time domain (FDTD) [8, 9], have demonstrated remarkable accuracy in simulating the interactions between electromagnetic waves and objects. However, when tasked with modeling complex and electrically-large targets, these methods encounter significant challenges such as the high computational resource demands and extensive time costs for matrix inversion $O(N^3)$, thereby revealing notable limitations in practical application. Researchers have been exploring methods to surmount these computational challenges, such as developing domain decomposition algorithms and acceleration techniques leveraging multi-CPU/GPU architectures [10–12]. These efforts have led to some improvements, but do not constitute a definitive solution. Innovative and potential solutions are still a pressing need to be explored.

In recent years, the rapid advancement of machine learning (ML) has introduced many innovative technical approaches for the modeling and analysis of target's RCS. For instance, artificial neural networks (ANNs) have been effectively utilized for the real-time prediction of 2D scattered fields [13] and the efficient computation of broadband monostatic RCS of morphing S-shaped cavities [14]. In addition, a hybrid model combining the autoregressive integrated moving average algorithm and the long short-term memory (LSTM) algorithm has been employed for the RCS sequence prediction [15]. Although these methods can model complex nonlinear relationships, they often require a large amount of data for training and can be prone to overfitting on small datasets.

Support vector regression (SVR), a robust machine learning model known for its exceptional nonlinear representation and generalization capabilities, and is less prone to overfitting on small datasets [16–18], has been increasingly incorporated into the field of CEM. Examples include modeling the target's RCS [19, 20] and the backscattering coefficient of the 3D sea surface [21]. To improve the RCS prediction accuracy of SVR, a physical

Submitted On: December 9, 2024
Accepted On: April 17, 2025

https://doi.org/10.13052/2025.ACES.J.400404

optics-inspired (POI) kernel function, which is a composite kernel function composed of the cosine function and the Gaussian kernel function, has been proposed [22]. Besides, a comprehensive analysis of the impact of various sampling schemes on the modeling of RCS using the POI kernel-based SVR (termed POI-SVR) has also been conducted. Results show that, compared to the centrally-located sampling (CLS), simple random sampling (SRS), and Latin hypercube sampling (LHS), uniform design (UD) and uniform design sampling (UDS) yield more representative training datasets, which can further improve the RCS modeling precision of POI-SVR. However, the cosine components in the POI kernel function are fixed without adjustable parameters, which limits the ability of POI-SVR to characterize the local fluctuation patterns of RCS. Moreover, the impact of data preprocessing on the modeling accuracy of SVR has not been explored yet.

In this paper, we introduce two angular frequency parameters into the cosine components of the POI kernel function, thereby augmenting the capacity of POI-SVR to characterize the local fluctuation characteristics of the target's RCS. Also, a physics-based data preprocessing method is proposed to further improve the accuracy of SVR in the modeling of the target's RCS. To facilitate the follow-up comparative analyses, the advanced POI-SVR presented herein is abbreviated as APOI-SVR.

The rest of this paper is organized as follows. Section II introduces the proposed APOI-SVR, including the improved POI kernel function, the physics-based data preprocessing method, and the training procedure of APOI-SVR. Several numerical validations for the proposed APOI-SVR are presented in section III. Finally, the conclusion is summarized in section IV.

II. THE PROPOSED APOI-SVR

APOI-SVR mainly improves the kernel function and the data preprocessing method on the basis of POI-SVR. Therefore, in this section, a detailed introduction is mainly given to the improved POI kernel function and the proposed data preprocessing method. The procedure of training APOI-SVR is illustrated at the end of this section as well.

A. Improved POI kernel

It is well accepted that kernel function is crucial for the performance of SVR. For linear problems, a linear kernel function is typically used; for periodic or quasi-periodic issues, a periodic kernel function is often chosen; and to enhance the local representation capability of SVR, a Gaussian kernel function is frequently selected. However, when it comes to complex problems, such as predicting the RCS of complex targets, a suitable kernel function needs to be carefully designed, which demands a deep understanding of the problem (i.e., prior knowledge).

In [22], the backscattered electric field $\boldsymbol{E}_{\mathrm{sca}}(\theta,\varphi,r)$ is approximated by:

$$\boldsymbol{E}_{\mathrm{sca}}(\theta,\varphi,r)=\mathrm{j}k\eta\frac{\exp(-jkr)}{4\pi r}\times\sum_{i=1}^{N}A_i\left[\hat{\boldsymbol{k}}\times(\hat{\boldsymbol{k}}\times\boldsymbol{J}_i')\right]\exp[-\mathrm{j}2\boldsymbol{k}\cdot\boldsymbol{r}_{i\vartheta}'],\quad(1)$$

where (θ,φ) is the incident angle of the electromagnetic wave, r is the distance from the origin to the observation point $\boldsymbol{r}$, $\boldsymbol{k}$ is the incident wave vector, $\hat{\boldsymbol{k}}=\boldsymbol{k}/k$, k and η represent the wave number and the wave impedance, $\boldsymbol{J}(\boldsymbol{r}')$ denotes the induced current at the source point $\boldsymbol{r}'$, A_i is the area of the facet S_i', and $\boldsymbol{r}_{i\vartheta}'$ is a point on the facet S_i'. Under far-field conditions, (1) can be further simplified to a function of (θ,φ):

$$\boldsymbol{E}_{\mathrm{sca}}(\theta,\varphi)=\sum_{i=1}^{N}a_i\boldsymbol{g}_i(\theta,\varphi),\quad(2)$$

with:

$$\begin{cases}a_i=\mathrm{j}k\eta A_i\exp(-\mathrm{j}kr)/4\pi r\\ \boldsymbol{g}_i(\theta,\varphi)=\left[\hat{\boldsymbol{k}}\times\left(\hat{\boldsymbol{k}}\times\boldsymbol{J}_i'\right)\right]\exp\left[-\mathrm{j}2\boldsymbol{k}\cdot\boldsymbol{r}_{i\vartheta}'\right]\end{cases},\quad(3)$$

where r is set to a sufficiently large constant (i.e., 10^9m). Notably, $\boldsymbol{g}_i(\theta,\varphi)$ contains the phase term:

$$\exp\left[-\mathrm{j}2\boldsymbol{k}\cdot\boldsymbol{r}_{i\vartheta}'\right]=\exp\left[-\mathrm{j}2k\left(\sin\theta\cos\varphi\cdot x_{i\vartheta}+\sin\theta\sin\varphi\cdot y_{i\vartheta}+\cos\theta\cdot z_{i\vartheta}\right)\right],\quad(4)$$

where $(x_{i\vartheta},y_{i\vartheta},z_{i\vartheta})$ are the coordinates of $\boldsymbol{r}_{i\vartheta}'$. The phase difference between facets introduces interference effects, leading to angular-dependent fluctuations of $\boldsymbol{E}_{\mathrm{sca}}(\theta,\varphi)$. To account for the fluctuation characteristics of the target's RCS and to maintain local representation capability of SVR, the following kernel function was proposed [22]:

$$K_P(x,x')=\prod_{i=1}^{2}\cos(x_i-x_i')\exp[-\frac{(x_i-x_i')^2}{2l_i^2}],\quad(5)$$

where $\boldsymbol{x}=(\theta,\varphi)$, l_1 and l_2 are scaling parameters. As inspired by physical optics (PO), the kernel function in (5) was termed POI kernel function. It should be noted that, the cosine components of $K_P(\boldsymbol{x},\boldsymbol{x}')$ are to characterize the fluctuation patterns of the target's RCS, the Gaussian components are to maintain good local representation capability.

However, for complex targets, their RCS responses often contain multiple harmonics due to the interactions among surfaces of varying sizes and orientations:

$$\sigma(\theta,\varphi)\approx\sigma_0+\sum_{m,n}C_{mn}\cos\left(\omega_{1m}\theta+\omega_{2n}\varphi+\psi_{mn}\right),\quad(6)$$

where σ_0 is the baseline RCS of the target, C_{mn} is the amplitude coefficient, ω_{1m} and ω_{2n} are angular frequencies, and ψ_{mn} is the initial phase offset of the $(m,n)-$th harmonic component. This decomposition highlights the multi-scale nature of RCS fluctuations. The cosine parts

of $K_P(\boldsymbol{x},\boldsymbol{x}')$ are fixed and devoid of adjustable parameters, assuming a single dominant frequency in the RCS spectrum. Thus, its capacity to accurately capture the various fluctuation patterns of RCS is inherently constrained. To tackle this issue, we introduce two angle frequency parameters (i.e.,ω_1 and ω_2), and proposed the following improved POI kernel function:

$$K'_P(\boldsymbol{x},\boldsymbol{x}') = \prod_{i=1}^{2} \cos[\omega_i(x_i - x'_i)] \exp[-\frac{(x_i - x'_i)^2}{2l_i^2}]. \quad (7)$$

It is clear that $K_P(\boldsymbol{x},\boldsymbol{x}')$ is a special case of $K'_P(\boldsymbol{x},\boldsymbol{x}')$ when $\omega_1 = \omega_2 = 1$. Figure 1 depicts the obvious variation in fluctuation patterns that arise from the introduction of the two angular frequency parameters ω_1 and ω_2. This demonstrates the potential of $K'_P(\boldsymbol{x},\boldsymbol{x}')$ in capturing the complex fluctuation patterns of the target's RCS.

To further illustrate the adaptability of the improved POI kernel function to the RCS patterns of complex targets, we analyzed the spectral properties of these two kernel functions.

First, consider the 1D case (θ-direction):

$$K'_p(\boldsymbol{x},\boldsymbol{x}') = \cos[\omega(x - x')] \exp\left(-\frac{(x - x')^2}{2l^2}\right). \quad (8)$$

The Fourier transform of this kernel is derived as follows:

$$\mathscr{F}\{K'_p\}(\omega') = \frac{1}{2}\left[\exp[-\frac{l^2(\omega' - \omega)^2}{2}] + \exp[-\frac{l^2(\omega' + \omega)^2}{2}]\right]. \quad (9)$$

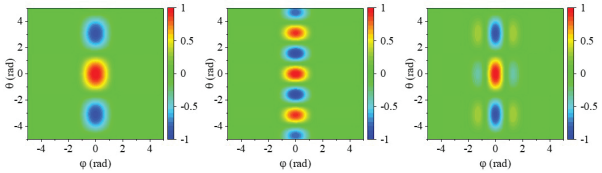
This result shows that the improved POI kernel acts as dual Gaussian bandpass filters centered at $\omega' = \pm\omega$, with bandwidth controlled by l.

For 2D RCS modeling, we have:

$$\mathscr{F}\{K'_p\}(\omega'_1, \omega'_2) \propto \prod_{i=1}^{2}\left[\begin{array}{l}\exp\left(-\frac{l_i^2(\omega'_i - \omega_i)^2}{2}\right) \\ +\exp\left(-\frac{l_i^2(\omega'_i + \omega_i)^2}{2}\right)\end{array}\right]. \quad (10)$$

This structure allows the improved POI kernel function to adaptively amplify frequency components near $(\pm\omega_1, \pm\omega_2)$. However, the POI kernel only amplifies frequency components near $(\pm 1, \pm 1)$.

The improved POI kernel function introduces angular frequency parameters ω_1 and ω_2, which can automatically adjust the center frequency of the filter according to the actual fluctuation frequency of the target's RCS (matching the multi-harmonic RCS spectrum). In this way, whether for simple or complex targets, the SVR model can find the most matching frequency to capture the RCS fluctuations. In contrast, the center frequency of the filter corresponding to the POI kernel function is fixed, easily lead to underfitting.



(a) (b) (c)

Fig. 1. Instances of the fluctuation patterns attributed to the varied angular frequency parameters ω_1 and ω_2. (a) $l_1 = 1, l_2 = 7, \omega_1 = \omega_2 = 1$, (b) $l_1 = 1, l_2 = 7, \omega_1 = 1, \omega_2 = 2$, (c) $l_1 = 1, l_2 = 7, \omega_1 = 2, \omega_2 = 1$.

B. Proposed data preprocessing method

In the IEEE dictionary of electrical and electronics terms, the definition of RCS is given by the following expression:

$$\sigma = \lim_{r\to\infty} 4\pi r^2 \frac{|\boldsymbol{E}_{\text{scat}}|^2}{|\boldsymbol{E}_{\text{inc}}|^2}, \quad (11)$$

where $\boldsymbol{E}_{\text{inc}}$ is the incident field. The unit of σ is m^2. Assuming $|\boldsymbol{E}_{\text{inc}}| = 1\,\text{V/m}$, then (11) can be simplified to:

$$\sigma = \lim_{r\to\infty} 4\pi r^2 |\boldsymbol{E}_{\text{scat}}|^2. \quad (12)$$

In the academic and industrial communities, it is also customary to use the following notation:

$$\sigma_{\text{dB}} = 10\lg(\sigma), \quad (13)$$

where the unit of σ_{dB} is dBsm.

It is noteworthy that $K'_P(\boldsymbol{x},\boldsymbol{x}')$ is derived from the scattered field formula presented in (2) and, thus, it is more apt for modeling the scattered field rather than modeling the target's RCS directly. Hence, we propose transforming the target's RCS into the magnitude of the scattered field $|\boldsymbol{E}_{\text{scat}}|$. Therefore, according to (12) and (13), we have:

$$|\boldsymbol{E}_{\text{scat}}| = \sqrt{\frac{\sigma}{4\pi r^2}} = \sqrt{\frac{10^{\sigma_{\text{dB}}/10}}{4\pi r^2}}. \quad (14)$$

Then, convert it into normalized scattered field strength:

$$\tilde{E}_{\text{scat}} = \frac{|\boldsymbol{E}_{\text{scat}}| - |\boldsymbol{E}_{\text{scat}}^{\min}|}{|\boldsymbol{E}_{\text{scat}}^{\max}| - |\boldsymbol{E}_{\text{scat}}^{\min}|}, \quad (15)$$

where $|\boldsymbol{E}_{\text{scat}}^{\max}|$ and $|\boldsymbol{E}_{\text{scat}}^{\max}|$ denote the maximum and the minimum scattered field strengths, respectively.

C. Training of APOI-SVR

This subsection primarily outlines the methodology for constructing APOI-SVR on previous derivations to model the target's RCS characteristics. The core focus is training APOI-SVR with sampled RCS data. Training APOI-SVR is essentially solving an optimization problem with constrains, that is:

$$\begin{aligned}\max_{\alpha,\hat{\alpha}} & \sum_{i=1}^{n} y_i(\alpha_i - \hat{\alpha}_i) - \varepsilon(\alpha_i + \hat{\alpha}_i) \\ & - \tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i - \hat{\alpha}_i)(\alpha_j - \hat{\alpha}_j)K'_P(\boldsymbol{x}_i,\boldsymbol{x}_j) \quad , \quad (16) \\ \text{s.t.} & \sum_{i=1}^{n}(\alpha_i - \hat{\alpha}_i) = 0, \quad \alpha_i, \hat{\alpha}_i \in [0, \mathscr{C}]\end{aligned}$$

where ε represents the width of the insensitive tube, and $\mathscr{C}$ is the penalty parameter. Assuming that n RCS data of the target, denoted as $\{(\theta_i, \varphi_i, \sigma_i)|i = 1, ..., n\}$, have been collected, the steps are as follows:

1. **Data preprocessing**: Apply (14) and (15) to preprocess the RCS data for training APOI-SVR.

2. **Hyperparameter optimization**: Adopt the Bayesian optimization method [23, 24] to obtain optimal values of hyperparameter l_1, l_2, ω_1, ω_2, and $\mathscr{C}$. In this work, an open-source Python library called "BayesO" is adopted.

3. **Model establishment**: Retrain APOI-SVR with these optimal hyperparameters to establish the approximation model for $\tilde{E}_{\text{scat}}(\theta, \varphi)$:

$$\tilde{E}_{\text{scat}}(\theta, \varphi) \approx \sum_{i=1}^{n} (\alpha_i - \hat{\alpha}_i) K_p'[(\theta, \varphi), (\theta_i, \varphi_i)] + b, \quad (17)$$

where α_i, $\hat{\alpha}_i$ and b are obtained by solving (16) with the sequential minimal optimization (SMO) [25].

4. **RCS modeling**: Based on (14), (15) and (17), establish the following surrogate model for modeling the target's RCS:

$$\sigma(\theta, \varphi) = 4\pi r^2 \times \left[\left| \boldsymbol{E}_{\text{scat}}^{\min} \right| + \left(|\boldsymbol{E}_{\text{scat}}^{\max}| - \left| \boldsymbol{E}_{\text{scat}}^{\min} \right| \right) \tilde{E}_{\text{scat}}(\theta, \varphi) \right]^2. \quad (18)$$

The overall flowchart of applying APOI-SVR for efficient modeling of the target's RCS is illustrated in Fig. 2. The improvements compared with POI-SVR [22] are marked with cyan.

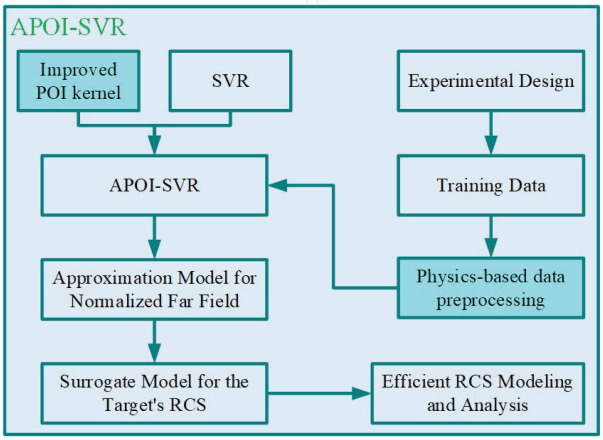
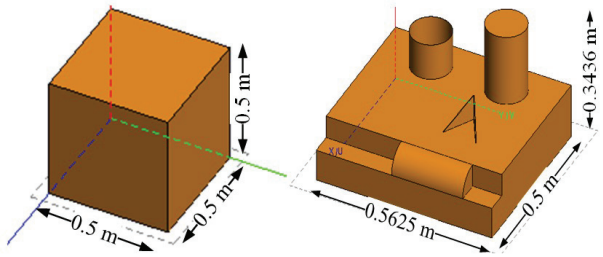


Fig. 2. Flowchart illustrating the application of APOI-SVR for efficient RCS modeling of a target.

III. VALIDATION RESULTS

In this section, we evaluate the performance of the proposed APOI-SVR across three critical dimensions: the accuracy of APOI-SVR in predicting the target's RCS, the efficiency of APOI-SVR in terms of model construction and prediction, and the applicability of APOI-SVR in rapidly modeling and analyzing the RCS of real-world targets.

Data preparation: To evaluate the performance of APOI-SVR, it is necessary to prepare some RCS data of targets. Due to the unknown RCS distribution of the target in practical applications, the uniform coverage assumption is rational. Informed by the analysis in [22], uniform design sampling (UDS) is based on this very assumption and achieves a smaller star discrepancy. Applying UDS can acquire more representative samples, thereby enhancing the RCS modeling accuracy of POI-SVR. Therefore, we employ UDS to sample the RCS data of both a simple Cube model and a complex SLICY model (widely used exemplary models in the field of CEM) in the upper space ($\theta \in [0, 0.5\pi]$, $\varphi \in [0, 2\pi]$), as shown in Fig. 3. The multilevel fast multipole algorithm (MLFMA) is utilized to compute the target's RCS at the angles in the UDS table, serving as the training data. The frequency of the incident electromagnetic waves is set to 1.0 GHz. Simple random sampling (SRS) is applied to guide the collection of test data. Details regarding the training/test datasets are presented in Table 1.



(a) (b)

Fig. 3. Two exemplary targets for the validation of the proposed APOI-SVR. (a) Conducting Cube with the side length of 0.5 m and (b) full-sized conducting SLICY model with dimensions of 0.5625 m × 0.5 m × 0.3436 m.

Table 1: Datasets information

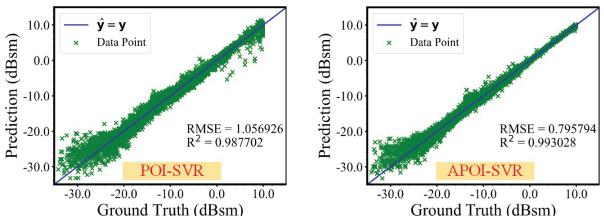
Target	Dataset Name	Number	Polarization
Cube	CVtrain	1296	VV
	CVtest	10000	VV
SLICY	SHtrain	1296	HH
	SHtest	10000	HH

A. Accuracy of APOI-SVR

In this subsection, we assess the modeling accuracy of APOI-SVR utilizing the training datasets CVtrain and

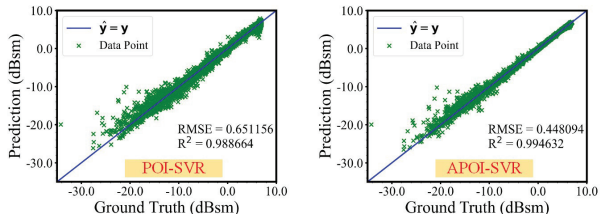
SHtrain, which correspond to the simple Cube model and the complex SLICY model, respectively. Equal training and hyperparameter optimization procedures are applied to both datasets, followed by numerical validations on the corresponding test datasets, CVtest and SHtest. The results are shown in Figs. 4 and 5, offering a comparative analysis between the RCS prediction accuracy of POI-SVR and APOI-SVR.

Figure 4 details the test results for the simple Cube model. Figure 4 (a) presents the performance of POI-SVR, with a root mean square error (RMSE) of 1.056926 and a coefficient of determination (R^2) of 0.987702. In contrast, Fig. 4 (b) displays the superior performance of APOI-SVR, with a reduced RMSE of 0.795794 and a higher R^2 of 0.993028. Figure 5 showcases the results for the complex SLICY model. Similar to the Cube model, APOI-SVR exhibits a higher precision in predicting the RCS of the complex SLICY model, as evidenced by a lower RMSE of 0.448094 and a higher R^2 of 0.994632 shown in Fig. 5 (b). As a competitor, POI-SVR obtains RMSE and R^2 of 0.651156 and 0.994097, respectively. Table 2 enumerates the RMSE reduction ratios of APOI-SVR compared with POI-SVR. Specifically, APOI-SVR reduces the predictive RMSE by 24.71% for the Cube model and by 31.18% for the SLICY model.



(a) (b)

Fig. 4. Test results for the simple Cube model: (a) POI-SVR and (b) APOI-SVR.



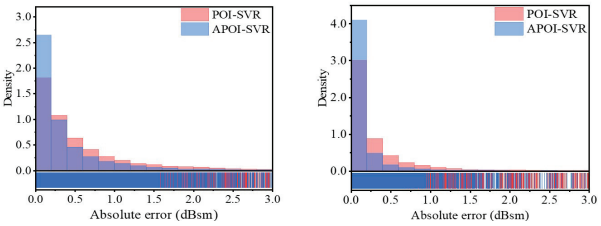
(a) (b)

Fig. 5. Test results for the complex SLICY model: (a) POI-SVR and (b) APOI-SVR.

Additionally, the absolute error distributions of POI-SVR and APOI-SVR are analyzed. Figure 6 presents the comparisons between the absolute error distributions of

Table 2: Predictive RMSE reduction ratios of APOI-SVR compared with POI-SVR

Target	Predictive RMSE		Reduction
	POI-SVR	APOI-SVR	
Cube	1.056926	0.795794	24.71%
SLICY	0.651156	0.448094	31.18%



(a) (b)

Fig. 6. Comparisons of absolute error distributions: (a) Cube and (b) SLICY.

POI-SVR and APOI-SVR across both the Cube model and the SLICY model. In the case of the Cube model, the absolute errors of APOI-SVR are primarily within the range of 0.0 to 2.0 dBsm, whereas POI-SVR displays a broader spread up to 3.0 dBsm. Similarly, in the case of the SLICY model, APOI-SVR maintains a tighter error range of 0.0 to 1.0 dBsm. demonstrating its robustness in handling complex targets.

On the whole, whether for the simple Cube model or the complex SLICY model, APOI-SVR achieves lower RCS prediction errors compared with POI-SVR. This indicates the effectiveness of the improved POI kernel function (see section IIA) and the physics-based data preprocessing method (see section IIB) in enhancing the RCS prediction accuracy of SVR, thereby offering a reliable and accurate modeling tool that facilitates the efficient analysis of the target's RCS characteristics.

B. Efficiency of APOI-SVR

In practical applications, the efficacy of an ML model is not solely appreciated by its prediction accuracy but also significantly by its implementation efficiency, including the time costs of optimizing hyperparameters with Bayesian optimization method and subsequent model retraining with the determined optimal hyperparameters, as well as the efficiency of the prediction process. Although the introduction of hyperparameters ω_1 and ω_2 into the improved POI kernel function significantly boosts the representation capability of APOI-SVR, it is essential to investigate any potential trade-offs in efficiency. Hence, a comprehensive analysis of both the implementation and prediction efficiency of APOI-SVR is warranted. It should be noted that, to ensure a compelling comparison, all time costs presented in this

paper are obtained under single-core operation with an Intel Core 2 Duo CPU T6670.

Figure 7 illustrates the time costs of APOI-SVR and POI-SVR, including implementation, prediction, and total times for both the Cube and SLICY models. It can be seen that in terms of implementation time, APOI-SVR takes less than POI-SVR. This reduction is attributed to the fact that, although the training time per iteration of APOI-SVR is slightly increased, the proposed physics-based data preprocessing method results in fewer optimization iterations necessary to reach convergence to the optimal solution (see Table 3). Consequently, the implementation time for APOI-SVR is shorter than that for POI-SVR. Regarding prediction time, the well-trained APOI-SVR and POI-SVR show negligible differences (less than 1.0 second). Finally, in terms of total time cost, APOI-SVR is less expensive than POI-SVR. Therefore, APOI-SVR surpasses POI-SVR in terms of overall efficiency.

C. Application of APOI-SVR

Having established the superior accuracy and efficiency of APOI-SVR in the previous subsections, we now apply APOI-SVR to analyze the RCS characteristics of a real-world target. Figure 8 depicts the aircraft model under consideration. Given its geometric symmetry, training data are sampled within the region where θ is in $[0°, 90°]$ and φ is in $[0°, 180°]$. In total, 1296 training samples are collected, preprocessed and subsequently used to train APOI-SVR. The well-trained APOI-SVR is ultimately applied to analyze the aircraft's RCS characteristics in the upper half space.

The frequency of the incident electromagnetic wave is set to 1.0 GHz. Figure 9 depicts the aircraft's RCS feature map, a 91×361 matrix, obtained by APOI-SVR. For comparison, the results acquired by MLFMA are given in Fig. 10. It can be observed that the results obtained by the two methods are highly consistent, indicating the reliability of APOI-SVR. However, it is worth noting that there is a significant difference in time cost (see Table 4). The total time cost of our APOI-SVR, containing data sampling, hyperparameter optimization, model training, and RCS prediction, is one-twelfth of that demanded by using MLFMA. This demonstrates the practicality of APOI-SVR for efficient RCS modeling of real-world targets.

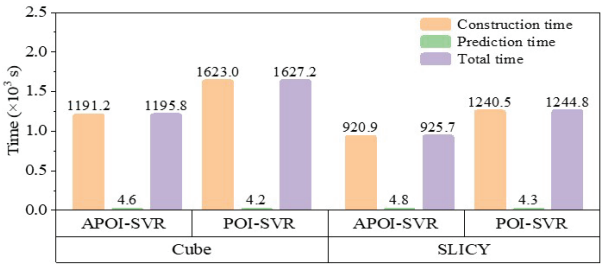


Fig. 7. Time cost comparison between POI-SVR and APOI-SVR.

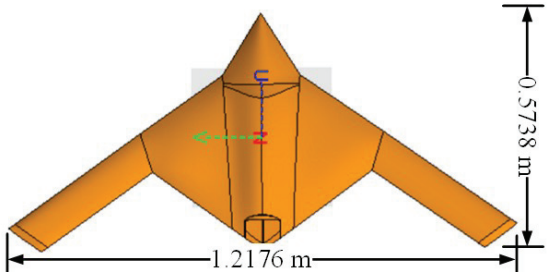


Fig. 8. Aircraft model for the illustration of efficient RCS modeling and analysis applying APOI-SVR.

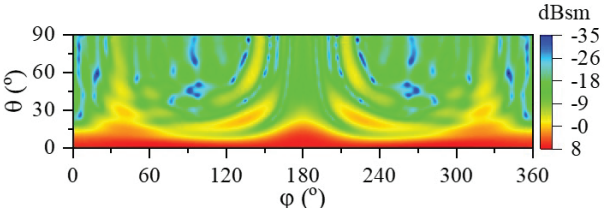


Fig. 9. RCS feature map of the aircraft obtained by APOI-SVR.

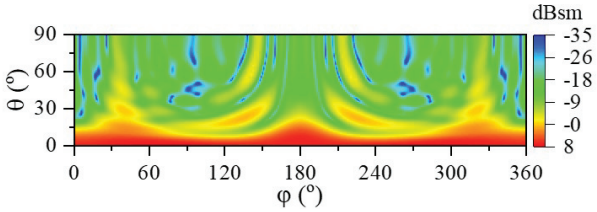


Fig. 10. RCS feature map of the aircraft acquired by MLFMA.

Table 3: Time costs and optimization iterations of hyperparameter optimization in the model implementation stage

Target	APOI-SVR			POI-SVR		
	Optimization Time Cost (s)	Optimization Iteration	Time Cost per Iteration (s)	Optimization Time Cost (s)	Optimization Iteration	Time Cost per Iteration (s)
Cube	1171.6	60	19.5	1617.6	100	16.2
SLICY	920.9	50	18.1	1224.7	80	15.3

Table 4: Time costs of acquiring the RCS feature map by APOI-SVR and MLFMA

Method	Sampling/Simulation Time Cost (s)	Optimization Time Cost (s)	Training Time Cost (s)	Prediction Time Cost (s)	Total Time Cost (s)
APOI-SVR	92738.9	1124.8	18.7	1.5	95509.6
MLFMA	1178628.5	—	—	—	1178628.5

IV. CONCLUSION

This paper proposes an advanced physical optics-inspired support vector regression (APOI-SVR) for the efficient modeling of a complex target's RCS. Two angular frequency parameters are introduced into the physical optics-inspired kernel function to address the various fluctuation patterns of RCS for complex targets, and a physics-based data preprocessing method is proposed to enable the model to efficiently learn the directly related physical quantity, i.e., the normalized electric field. Numerical validations conducted on both a simple Cube model and a complex SLICY model have confirmed that, compared with POI-SVR, the new-proposed APOI-SVR effectively reduces the RMSE in RCS prediction by over 24.7%. Moreover, it maintains high predictive efficiency, capable of completing the prediction of 10,000 samples in the test dataset within 5.0 seconds. Notably, although the introduction of two angular frequency parameters slightly increases the training time for each iteration in the hyperparameter optimization process, the proposed physics-based data preprocessing method reduces the required number of optimization iterations. As a result, in terms of overall efficiency, APOI-SVR outperforms POI-SVR.

Additionally, the application of APOI-SVR to an aircraft model has illustrated its practical efficacy in generating RCS feature maps with high precision and efficiency compared to the well-known MLFMA. This practical application indicates that APOI-SVR may be a valuable tool in the field of electromagnetic scattering analysis. Future research will commit to expanding the applicability of APOI-SVR, including the enhancements tailored for complex targets with coatings.

ACKNOWLEDGMENT

This work is based on the research supported in part by the Postdoctoral Fellowship Program (Grade C) of China Postdoctoral Science Foundation under Grant Number GZC20241307, and in part by the National Natural Science Foundation of China under Grant 62231021, Grant 62071347, and Grant 62171335. Corresponding author: Donghai Xiao.

REFERENCES

[1] M. I. Skolnik, *Radar Handbook*, 3rd ed. New York, NY: McGraw-Hill, 2008.

[2] E. F. Knott, J. F. Schaeffer, and M. T. Tulley, *Radar Cross Section*, 2nd ed. Rayleigh, NC: SciTech, 2004.

[3] M. Cavagnaro, E. Pittella, and S. Pisa, "Numerical evaluation of the radar cross section of human breathing models," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 30, no. 12, pp. 1354-1359, Aug. 2015.

[4] R. Haupt, S. Haupt, and D. Aten, "Evaluating the radar cross section of maritime radar reflectors using computational electromagnetics," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 24, no. 4, pp. 403-406, Aug. 2009.

[5] M. M. Sheikh, "Analysis of reader orientation on detection performance of Hilbert curve-based fractal chipless RFID tags," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 39, no. 6, pp. 490-504, June 2024.

[6] J.-M. Jin, *The Finite Element Method in Electromagnetics*. Hoboken, NJ: Wiley-IEEE Press, 2014.

[7] C. Qin, X. Wang, and N. Zhao, "EMFEM: A parallel 3D modeling code for frequency-domain electromagnetic method using goal-oriented adaptive finite element method," *Comput. Geosci.*, vol. 178, Art. no. 105403, Sep. 2023.

[8] F. L. Teixeira, C. Sarris, Y. Zhang, D.-Y. Na, J.-P. Berenger, Y. Su, M. Okoniewski, W. C. Chew, V. Backman, and J. J. Simpson, "Finite-difference time-domain methods," *Nature Rev. Methods Primers*, vol. 3, p. 75, 2023.

[9] A. Taflove and S. C. Hagness, *Computational Electromagnetics: The Finite-Difference Time-Domain Method*, 2nd ed. Norwood, MA: Artech House, 2000.

[10] P. D. Cannon and F. Honary, "A GPU-accelerated finite-difference time-domain scheme for electromagnetic wave interaction with plasma," *IEEE Trans. Antennas Propagat.*, vol. 63, no. 7, pp. 3042-3054, July 2015.

[11] T. Maceina, P. Bettini, G. Manduchi, and M. Passarotto, "Fast and efficient algorithms for computational electromagnetics on GPU architecture," *IEEE Trans. Nucl. Sci.*, vol. 64, no. 7, pp. 1983-1987, July 2017.

[12] Y. Xu, H. Ma, and R. Jiang, "Collaborating CPU and GPU for the electromagnetic simulations with the FDTD algorithm," *Concurrency Com-*

put., Pract. Exper., vol. 29, no. 4, p. e3859, Feb. 2017.

[13] W.-W. Zhang, D.-H. Kong, X.-Y. He, and M.-Y. Xia, "A machine learning method for 2-D scattered far-field prediction based on wave coefficients," *IEEE Antennas Wireless Propag. Lett.*, vol. 22, no. 5, pp. 1174-1178, May 2023.

[14] R. Weng, D. Sun, W. Yang, X. Chen, and W. Lu, "Efficient broadband monostatic RCS computation of morphing S-Shape cavity using artificial neural networks," *IEEE Antennas Wireless Propag. Lett.*, vol. 22, no. 2, pp. 263-267, Feb. 2022.

[15] J. Wu, C. Liu, L. Yang, and L. Tang, "Research on RCS sequence prediction based on ARIMA-TPA-LSTM algorithm," in *2023 China Automation Congress (CAC)*, Chongqing, China, pp. 1887-1892, Nov. 2023.

[16] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, pp. 155-161, 1997.

[17] D. Basak, S. Pal, and D. C. Patranabis, "Support vector regression," in *Proc. Neural Inf. Process. Lett. Rev.*, vol. 11, no. 10, pp. 203-224, 2007.

[18] M. Awad and R. Khanna, "Support vector regression," in *Efficient Learning Machines,* Mariette Awad and Rahul Khanna, Eds. Berlin, Germany: Springer, 2015, pp. 67-80.

[19] Z. Zhang, P. Wang, and M. He, "Wideband monostatic RCS prediction of complex objects using support vector regression and grey-wolf optimizer," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 38, no.8, pp. 609-615, Aug. 2023.

[20] Z. Zhang and M. He, "Fast prediction of electromagnetic scattering fields based on machine learning and PSO algorithm," in *2022 IEEE 10th Asia-Pacific Conference on Antennas and Propagation (APCAP)*, Xiamen, China, pp. 1-2, 2022.

[21] C. Dong, X. Meng, L. Guo, and J. Hu, "3D sea surface electromagnetic scattering prediction model based on IPSO-SVR," *Remote Sens.*, vol. 14, no. 18, p. 4657, Sep. 2022.

[22] D. Xiao, L. Guo, W. Liu, and M. Hou, "Efficient RCS prediction of the conducting target based on physics-inspired machine learning and experimental design," *IEEE Trans. Antennas Propag.*, vol. 69, no. 4, pp. 2274-2289, Apr. 2021.

[23] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, pp. 2951-2959, 2012.

[24] J. Kim and S. Choi, "BayesO: A Bayesian optimization framework in Python," *J. Open Source Softw.*, vol. 8, no. 90, p. 5320, Oct. 2023.

[25] G. W. Flake and S. Lawrence, "Efficient SVM regression training with SMO," *Mach. Learn.*, vol. 46, nos. 1-3, pp. 271-290, 2002.

Chenge Shi received her M.S. and Ph.D. degrees in Radio Physics from Xidian University, Xi'an, Shaanxi, China, in 2017 and 2023, respectively. She is currently a Researcher with AVIC Xi'an Aircraft Industry Group Company Ltd., primarily interested in aircraft manufacturing and testing.

Rui Cai received his B.S. degree from Xidian University in 2011, and has been enthusiastically contributing to AVIC Xi'an Aircraft Industry Group Company Ltd. He is a Senior Engineer, mainly engaged in the demonstration of test scheme and system simulation.

Wei Dong is the Chief Engineer of AVIC Xi'an Aircraft Industry Group Company Ltd. He is mainly engaged in the testing of modern airborne systems, which include electromechanical components, avionics, and flight control systems.

Donghai Xiao received the B.S. degree in Applied Physics and the Ph.D. degree in Radio Physics from Xidian University, Xi'an, China, in 2014 and 2023, respectively. He is currently a Postdoctoral Researcher with Hangzhou Institute of Technology, Xidian University, Hangzhou, China. His research interests include the areas of computational electromagnetics, radar signal processing and machine learning.

Compressing Electromagnetic Field by Rational Interpolation of the Spherical Wave Expansion Coefficients

Haobo Yuan[1], Pu Yang[1], Nannan Wang[1], Yi Ren[1], and Shasha Li[2]

[1]School of Electronic Engineering
Xidian University, Xi'an, Shaanxi, China
useryuanhaobo@163.com, 3024396382@qq.com, wnannan2001@163.com, yren@xidian.edu.cn

[2]Beijing Institute of Space Long March Vehicle
Beijing, China
heiyu-kuaidou@hotmail.com

Abstract – It is of great significance to obtain the electromagnetic field radiated by an antenna or scattered by an object over a frequency band. But this data often occupies so large a memory that cannot be applied readily. This paper proposes to compress the field based on the spherical harmonic transformation (SHT) and rational interpolation. First, the tangential electric field over a sphere surrounding the antenna is obtained by simulation or measurement. Then, this field is converted into the spherical harmonic coefficients, which are sparse discrete spectra. Finally, these coefficients are interpolated over the whole frequency band with only a few sampling points. Numerical examples show that the proposed algorithm can compress the data of the near field of a rectangular waveguide antenna by about 17278 times, and those of the far field scattered from an UAV by about 103 times.

Index Terms – Antenna, data compression, rational interpolation, spherical harmonic transform.

I. INTRODUCTION

The data of the electromagnetic field radiated by an antenna or scattered from an object is widely applied in many engineering scenarios, such as radar imaging, antenna measurement and base station deployment. As shown in Fig. 1, the field often varies rapidly as frequency and scanning angle changes. Therefore, one needs to sample quite densely in both frequency and angle to accurately represent the field. Many compression techniques have been developed to reduce this large amount of data.

The method introduced by Burnside et al. is based on radar imaging technology [1-3]. It extracted the scattering centers according to the radar images. The scattering field from an individual scattering center can be expressed as a complex exponential function of frequency and angle. As a result, the radar cross section (RCS) is compressed drastically. There are some additional methods based on the theory of scattering centers, such as the matrix pencil method [4] and the CLEAN algorithm [5, 6].

Regarding the data of the electromagnetic field as a matrix, one can make use of the well-established image compression algorithms. These algorithms often exploit the low-rank property of a particular matrix. In order to compress the near field data, Wu et al. proposed the CUR decomposition [7] and Zhao et al. proposed the skeletonization-scheme [8]. Guo et al. applied the butterfly scheme [9] to compress the system matrix generated by the combined-field integral equation. The most widespread ones are the threshold discrete Fourier transform (TDFT) method [10, 11] and the truncated singular value decomposition (SVD) method [12, 13].

Another type of compression method applies compressive sensing (CS) [14–19] to reconstruct the antenna radiation pattern in the antenna measurements. The CS process often uses a transform, such as discrete cosine transform or discrete Fourier transform, that renders the data of the field to be a sparse vector in the transform domain. Minimizing the l_1-norm of this vector will reconstruct the field data with much fewer random measurements. This method sheds some light on the proposed method, which transforms the electromagnetic fields into the spherical harmonic spectra in Fig. 2 by spherical harmonic transformation (SHT). Fortunately, the spherical harmonic spectra are low-pass, discrete and sparse. Furthermore, the l_1-norm minimization in CS may be time-consuming if the iterative procedure diverges, whereas SHT is a deterministic algorithm with $O(N^3)$ operations, where N is the truncation order.

Some use SHT to compress the pattern of an antenna. Reference [20] expanded the field by SHT and Slepian decomposition. Reference [21] applied the

Submitted On: August 4, 2024
Accepted On: April 10, 2025

https://doi.org/10.13052/2025.ACES.J.400405

sparse spherical harmonic expansion with compressed sensing to expand the far field of an antenna. None of them discussed the frequency-sweeping technique, thus the compression ratio for a frequency band will be limited. Reference [22] applied the 2-D scalar SHT and the windowed interpolation for compression. As is discussed in [22], this commonly used rational interpolation method is numerically unstable as the number of sampling points increases. Therefore, the interpolated function may have some spikes in the curve, called "Froissart doublets" [23].

Since the data of the field is also a function of frequency, we need to implement frequency sweeping efficiently. Reference [24] considered the field radiated by an antenna at a specific angle as a function of frequency, which is expanded by the Chebyshev polynomials. Actually, most of the physical quantities in an electronic system are not polynomials, but rational polynomials. Therefore, this paper applies the rational interpolation based on the Loewner matrix as the frequency sweeping tool.

In summary, the proposed method combines two techniques to compress the data of field in Fig. 1. On the one hand, the field at a specific frequency is converted into spherical harmonic spectra in Fig. 2. On the other hand, the spherical harmonic spectra over a frequency band are approximated by the rational interpolation, which requires the spectra at only a few sampling frequencies.

This paper is organized as follows. Section II introduces the theory of SHT and rational interpolation, and gives the flowchart of the proposed algorithm. Section III validates the proposed algorithm with two numerical examples. Finally, conclusions are drawn in section IV.

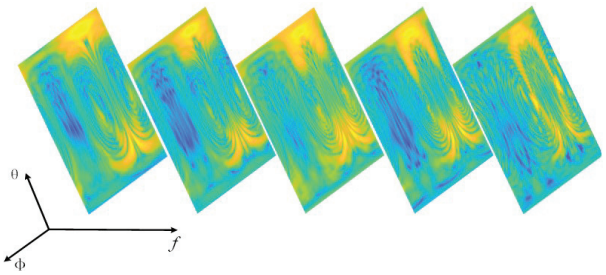


Fig. 1. Electric field varies with frequency.

Fig. 2. Spherical harmonic spectra vary with frequency.

II. FORMULATION OF THE ALGORITHM

A. Spherical harmonic transform

Spherical harmonic transform is a well-established algorithm in near-field antenna measurement. The time-harmonic electromagnetic field in a source free space generated by the antenna can be expanded by the spherical harmonics as in [25]:

$$E = -\sum_{n=1}^{N}\sum_{m=-n}^{n}(a_{n,m}M_{n,m} + b_{n,m}N_{n,m}), \tag{1}$$

$$H = \frac{-1}{120\pi}\sum_{n=1}^{N}\sum_{m=-n}^{n}(a_{n,m}N_{n,m} + b_{n,m}M_{n,m}), \tag{2}$$

where N is the truncation order of the spherical harmonics, $a_{n,m}$ and $b_{n,m}$ are the spherical harmonic coefficients and the range of subscripts is $1 \le n \le N$, $-n \le m \le n$. The vector spherical harmonics in the spherical coordinate system are:

$$M_{n,m} = C_{n,m}\cdot\begin{bmatrix}\frac{jm}{\sin\theta}Z_n(kr)P_n^{|m|}(\cos\theta)\,\hat{\theta}\\ -Z_n(kr)\frac{d}{d\theta}P_n^{|m|}(\cos\theta)\,\hat{\varphi}\end{bmatrix}\cdot e^{jm\varphi}, \tag{3}$$

$$N_{n,m} = C_{n,m}\cdot\begin{Bmatrix}\frac{Z_n(kr)}{kr}n(n+1)P_n^{|m|}(\cos\theta)\,\hat{r}\\ +W_n(kr)\begin{bmatrix}\frac{d}{d\theta}P_n^{|m|}(\cos\theta)\,\hat{\theta}\\ +\frac{jm}{\sin\theta}P_n^{|m|}(\cos\theta)\,\hat{\varphi}\end{bmatrix}\end{Bmatrix}\cdot e^{jm\varphi}, \tag{4}$$

where θ and φ are the elevation angle and the azimuth angle respectively, n is the degree, m is the order, $P_n^{|m|}(\cos\theta)$ is the associated Legendre function, $Z_n(kr)$ is the spherical Bessel function, k is the wave number of free space, $C_{n,m} = (-1)^m\sqrt{\frac{(2n+1)(n-|m|)!}{4\pi n(n+1)(n+|m|)!}}$, $W_n(kr) = \frac{1}{kr}\frac{d}{dr}[rZ_n(kr)]$.

In the spherical near-field antenna measurement, we acquire the tangential near electric field on a sphere surrounding the antenna by a mechanical scanning procedure:

$$E^{\tan} = E_\theta\hat{\theta} + E_\varphi\hat{\varphi}, \tag{5}$$

where E_θ and E_φ are the components of $E^{\tan}$ in the θ and φ direction. The unknown spherical harmonic coefficients can be evaluated by the following integrals:

$$a_{n,m} = -\frac{C_{n,m}}{Z_n(kr_0)}\begin{Bmatrix}\int_0^\pi\left[\int_0^{2\pi}E_\theta(r_0,\theta,\varphi)e^{-jm\varphi}d\varphi\right]\\ \cdot jmP_n^{|m|}(\cos\theta)\,d\theta\\ +\int_0^\pi\left[\int_0^{2\pi}E_\varphi(r_0,\theta,\varphi)e^{-jm\varphi}d\varphi\right]\\ \cdot\frac{d}{d\theta}\left[P_n^{|m|}(cos\theta)\right]\sin\theta\,d\theta\end{Bmatrix}. \tag{6}$$

$$b_{n,m} = -\frac{C_{n,m}}{W_n(kr_0)}\begin{Bmatrix}\int_0^\pi\left[\int_0^{2\pi}E_\theta(r_0,\theta,\varphi)e^{-jm\varphi}d\varphi\right]\\ \cdot\frac{d}{d\theta}\left[P_n^{|m|}(\cos\theta)\right]\sin\theta\,d\theta\\ +\int_0^\pi\left[\int_0^{2\pi}E_\varphi(r_0,\theta,\varphi)e^{-jm\varphi}d\varphi\right]\\ \cdot jmP_n^{|m|}(\cos\theta)\,d\theta\end{Bmatrix}. \tag{7}$$

The above double integral consists of an inner integral and an outer integral. The inner integral is just a

Fourier integral respect to , which can be evaluated efficiently by the Fast Fourier Transform (FFT). The outer integral is usually calculated by a Gaussian Quadrature. Given the near fields of the antenna, acquiring the spherical harmonic coefficients by (6) and (7) is referred to as the forward SHT. Given the spherical harmonic coefficients of the antenna, evaluating the near fields by (1) is called the inverse SHT. Owing to the FFT, both forward SHT and inverse SHT have the computational complexity of $O(N^3)$, which can be reduced to $O(N^2 \log N)$ by a novel fast SHT [26].

The expansion coefficients $a_{n,m}$ and $b_{n,m}$ are also termed the spherical harmonic spectra of the electromagnetic field. These are low-pass and discrete spectra which can be stored easily. Because $1 \le n \le N$, $-n \le m \le n$, the number of coefficients $a_{n,m}$ is $N \times (N+2)$, and it also applies for $b_{n,m}$. According to [26], we often take $N = kd + 10$, where d is the size of the antenna; the memory requirement for this spectrum is trivial.

Furthermore, among all the $2N \times (N+2)$ coefficients, only a small portion are relatively large quantities, whereas the rest are so small they can be neglected. Therefore, if we only store the non-zero coefficients, then the memory requirement could be reduced significantly. The sparsity pattern of these coefficients is shown in Figs. 12 and 13.

Obviously, (6) and (7) are applied to represent the near field at a single frequency. Often we need the field over a wide frequency band, which will be addressed by a frequency sweeping algorithm called rational interpolation based on Loewner matrix [27, 28].

B. Rational interpolation respect to frequency

Consider only one of the coefficients $a_{n,m}$ and $b_{n,m}$. This coefficient is denoted by $x(s)$, where s represents the frequency. Suppose $2p-1$ sampling data have been obtained by the forward SHT for the near field of an antenna:

$$x(s_i) = x_i, \ i = 1, 2, 3, \cdots, 2p-1, \tag{8}$$

where $s_1 < s_2 < \cdots < s_{2p-1}$ are the sampling frequenci-es. We partition these data into two groups:

$$x(\lambda_i) = w_i, i = 1, 2, 3, \cdots, p, \tag{9}$$

$$x(\mu_j) = v_j, j = 1, 2, 3, \cdots, p-1. \tag{10}$$

With (9), $x(s)$ can be expressed by the following rational approximation in barycentric form:

$$x_L(s) = \sum_{i=1}^{p} \frac{\alpha_i w_i}{s - \lambda_i} \Big/ \sum_{i=1}^{p} \frac{\alpha_i}{s - \lambda_i}, \tag{11}$$

where $\alpha_i \, (i = 1, \cdots, p)$ are unknown coefficients to be determined by (10). Evaluating $x_L(s)$ at the points in the second partition (10) leads to:

$$x_L(\mu_j) = \sum_{i=1}^{p} \frac{\alpha_i w_i}{\mu_j - \lambda_i} \Big/ \sum_{i=1}^{p} \frac{\alpha_i}{\mu_j - \lambda_i} = v_j. \tag{12}$$

Subsequently, we have:

$$\sum_{i=1}^{p} \frac{v_j - w_i}{\mu_j - \lambda_i} \alpha_i = 0, \tag{13}$$

which is written in compact matrix form as:

$$\begin{bmatrix} \frac{v_1 - w_1}{\mu_1 - \lambda_1} & \cdots & \frac{v_1 - w_p}{\mu_1 - \lambda_p} \\ \vdots & \frac{v_j - w_i}{\mu_j - \lambda_i} & \vdots \\ \frac{v_{p-1} - w_1}{\mu_{p-1} - \lambda_1} & \cdots & \frac{v_{p-1} - w_p}{\mu_{p-1} - \lambda_p} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix} = 0. \tag{14}$$

The system matrix on the left side of (14) is the so-called Loewner matrix based on the adopted partition of the samples, and the unknown coefficients $\alpha_i$can be readily evaluated by the SVD of the Loewner matrix. Then, the rational polynomial (11) goes through all the $2p-1$ points, and can be viewed as an approximation of the unknown function $x(s)$.

C. Compressing the field of an antenna or scatterer

The above rational interpolation is suitable for the scalar function, and can be generalized to interpolate a vector function, such as the vector containing all the coefficients $a_{n,m}$ and $b_{n,m}$, for a frequency band $s \in [s_a, s_b]$. More details are given in [3].

The flowchart Fig. 3 presents the algorithm to compress the field of an antenna. It mainly includes four components. The first one is to obtain the tangential near electric field over a sphere surrounding the antenna by

ALGORITHM 1:

1. Start from a given error tolerance τ and three frequencies: $s_1 = s_a$, $s_2 = 0.5(s_a + s_b)$, $s_3 = s_b$. Set $p = 2$.
2. Compute or measure the tangential electric field $\boldsymbol{E}^{\tan}$ over a sphere surrounding the antenna at all the frequencies $s_1, s_2, \ldots s_{2p-1}$.
3. Convert the near-field $\boldsymbol{E}^{\tan}$ into coefficients $a_{n,m}(s_i)$ and $b_{n,m}(s_i)$ by forward SHT in (6) and (7).
4. Interpolate the non-zero coefficients $a_{n,m}(s)$, $b_{n,m}(s)$ at all the $(2p-1)$ sampling frequencies by rational interpolation.
5. Estimate the error ε_{est}^{i} of the i^{th} segments $[s_i, s_{i+1}]$ by the technique in [3], where $i = 1, 2, \ldots, 2p-2$.
6. Find the segment I and II with the first and the second largest errors ε_{est}^{I} and ε_{est}^{II}, respectively.
7. If $\varepsilon_{est}^{I} > \tau$, add two frequencies: $s_{2p} = 0.5(s_I + s_{I+1})$, $s_{2p+1} = 0.5(s_{II} + s_{II+1})$, set $p = p + 1$, and go to 2. If $\varepsilon_{est}^{I} \le \tau$, the rational interpolation converges, go to 8.
8. To retrieve the near field at an arbitrary frequency s_x, compute the coefficients $a_{n,m}(s_x), b_{n,m}(s_x)$ by (11), and convert them into the near-field over any enclosing sphere by inverse SHT in (1).

Fig. 3. Flowchart of the compression algorithm.

computational electromagnetic algorithms or measurements. The second one is the forward SHT in part A. The third one is the rational interpolation in part B. And the last one is the inverse SHT in part A. All together, we can compress the field into the spherical harmonic coefficients and retrieve the field at any frequency efficiently.

III. NUMERICAL RESULTS

The proposed method is validated with the field radiated by a waveguide antenna and the field scattered from an UAV. As the frequency changes, the field of the former varies slightly, while those of the latter varies rapidly. Therefore, we investigate the former over a wide band and the latter over a narrow band. In order to evaluate the accuracy of the compression methods, we define the relative error as:

$$\varepsilon = \sqrt{\frac{\sum_{j=1}^{M}\sum_{i=1}^{N}\left|E_c(\theta_i,\varphi_j)-E_r(\theta_i,\varphi_j)\right|^2}{\sum_{j=1}^{M}\sum_{i=1}^{N}\left|E_r(\theta_i,\varphi_j)\right|^2}}, \qquad (15)$$

where θ_i, φ_j are the sampling points along the two angles in the spherical coordinate, M and N are the corresponding number of points, E_c represents the electric field obtained by a compression method, and E_r represents the reference field, which is obtained by MoM directly. Furthermore, the definition of compression ratio (CR) is [28]:

$$CR = \frac{size\ of\ the\ original\ data}{size\ of\ the\ compression\ data}. \qquad (16)$$

A. Rectangular waveguide antenna

Figure 4 shows the structure of a WR187 rectangular waveguide antenna, which has an aperture of 47.55×22.15 mm, and is 180 mm in length. This antenna operates at C-band between 4 GHz and 6 GHz. The proposed algorithm is implemented to compress and restore the near field. It is pointed out that the tangential electric field $E^{\tan}$ over the enclosing sphere is obtained by MoM, which is the second step of the algorithm. The radius of the sphere is 150 mm, and the truncation order in SHT is N= 14.

Subsequently, the near field is converted into spherical wave expansion coefficients. Among all the 448 coefficients, there are only 102 non-zero ones, as shown in Figs. 5 and 6. These coefficients are fitted by the rational interpolation with only 11 frequencies and the tolerance in Fig. 3 is $\tau = 0.03$. In other words, the MoM simulation is implemented 11 times. Figures 7 and 8 show that the real and imaginary parts of the interpolated $a_{1,1}$ and $b_{14,7}$ are almost identical to the reference. Therefore, the spherical wave expansion coefficients at an arbitrary frequency in the band can be predicted by (10). For example, we compute the coefficients at 5.33 GHz, then restore the near electric field on the enclosing sphere by inverse SHT. Figures 9 and 10 show that the near-field is in good agreement with the reference result.

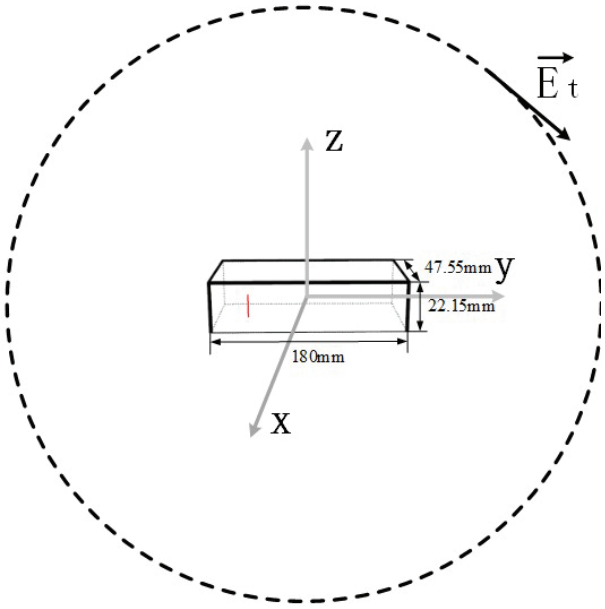


Fig. 4. Waveguide antenna and enclosing sphere.

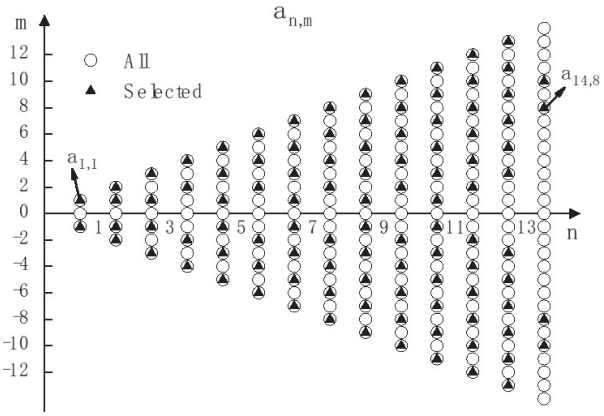


Fig. 5. Sparsity of spherical harmonic spectra $a_{n,m}$.

Finally, the compression ratio of the near-field is considered. First, we evaluate the memory requirement of the near field without compression. Suppose there are 200 uniformly distributed frequencies from 4 GHz to 6 GHz, and the field at each frequency has 360 azimuthal points and 180 elevational points on the surrounding sphere. Then the memory requirement of the near-field complex vectors is 200×360×180×3×16≈622 MB. Thus, the proposed method needs only 11 frequencies over the whole band, and each frequency has only 102 non-zero spherical wave expansion coefficients. The memory requirement is 11×102×2×16≈0.036 MB, which is negligible. Therefore, the compression ratio is about 17278.

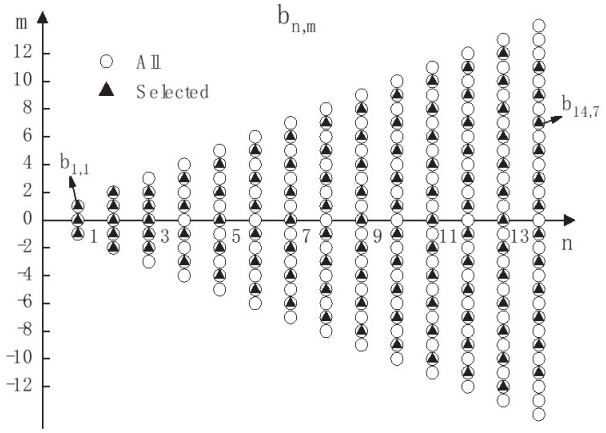


Fig. 6. Sparsity of spherical harmonic spectra $b_{n,m}$.

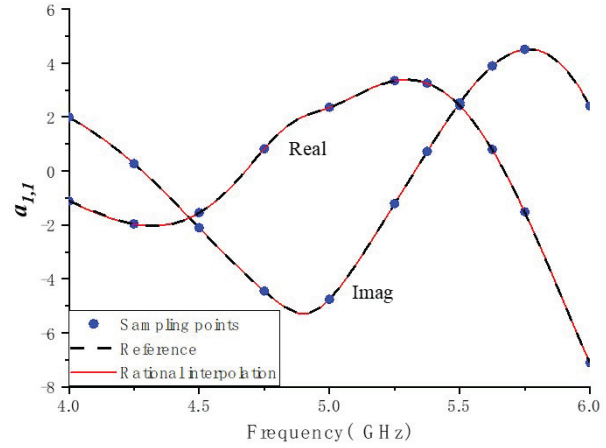


Fig. 7. Rational interpolation of the coefficient $a_{1,1}$.

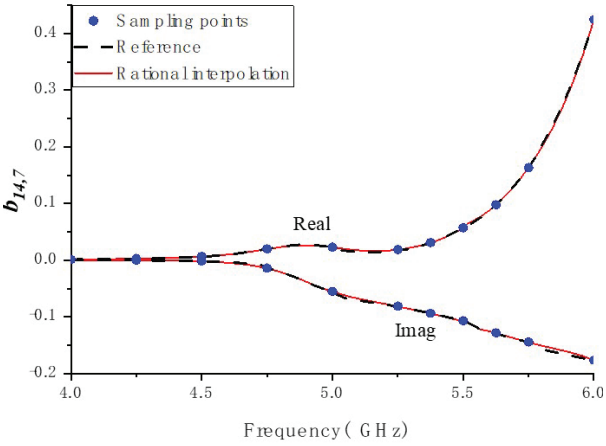


Fig. 8. Rational interpolation of the coefficient $b_{14,7}$.

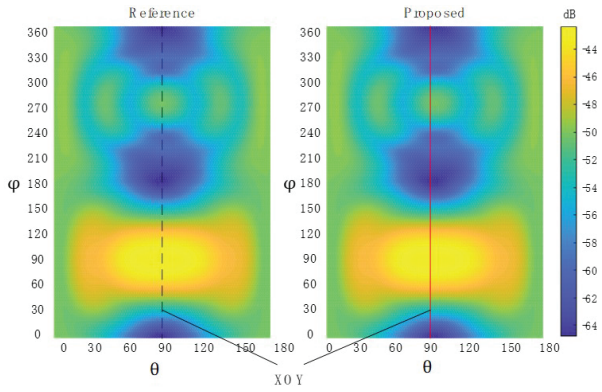


Fig. 9. Near-field of waveguide antenna at 5.33 GHz.

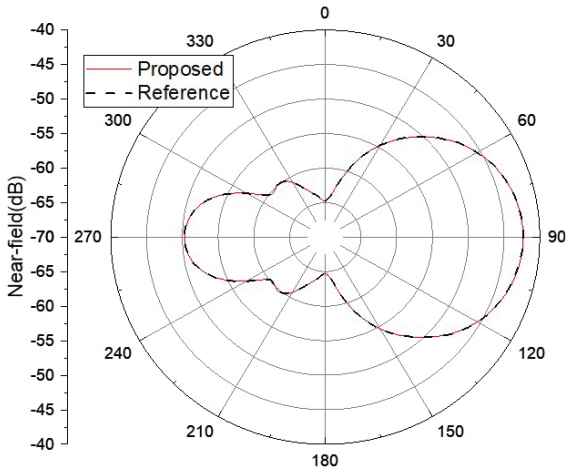


Fig. 10. Near-field of waveguide antenna in XOY cut at 5.33 GHz.

B. UAV RCS

Figure 11 shows the model of a UAV, which is about 5 m in length. The UAV is illuminated by a plane wave $E^{\text{inc}}=\hat{z}e^{-jky}$, and the operating frequency band is from 1180 MHz to 1220 MHz. The proposed algorithm will be applied to compress the bistatic RCS of the UAV. Similarly, the second step of the algorithm is computing the tangential near-field $E^{\tan}$ of the UAV over the enclosing sphere by MoM. The radius of the sphere is 5 m, and the truncation order in SHT is $N=139$.

Then, the near-field is converted into spherical wave expansion coefficients. There are 15761 non-zero coefficients among all the 39198 coefficients, as shown in Figs. 12 and 13. These non-zero coefficients are interpolated over the frequency band with only 19 frequencies. Figures 14 and 15 show that the real parts and imaginary parts of the interpolated $a_{125,0}$ and $b_{2,-1}$are in good agreement with those of the reference. Also, we compute the coefficients at an arbitrary frequency, say 1216 MHz, by (10), and then restore the far-field or RCS by inverse

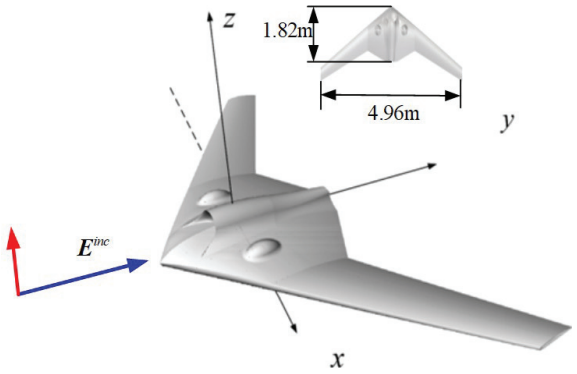


Fig. 11. Structure of a UAV.

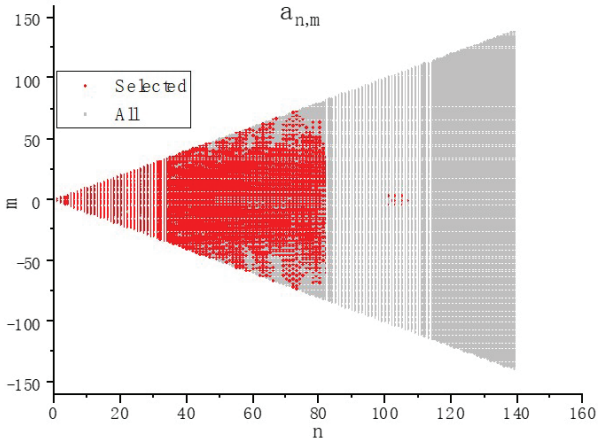


Fig. 12. Sparsity of spherical harmonic spectra $a_{n,m}$.

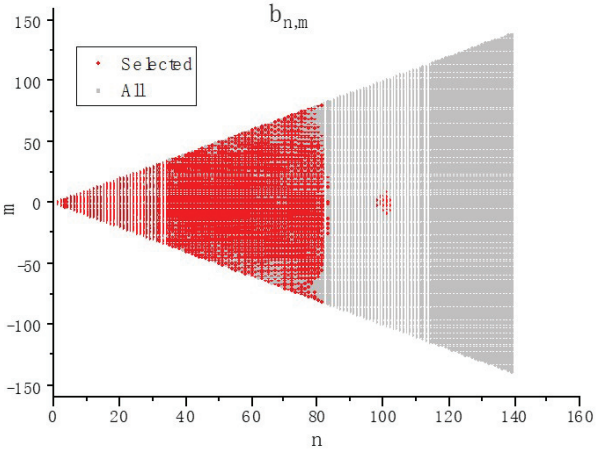


Fig. 13. Sparsity of spherical harmonic spectra $b_{n,m}$.

SHT. Figures 16 and 17 show that the restored far-field is almost the same as the reference result.

Then, the compression ratio is considered. First, we evaluate the memory requirement of the RCS without compression. Suppose there are 40 uniformly distributed

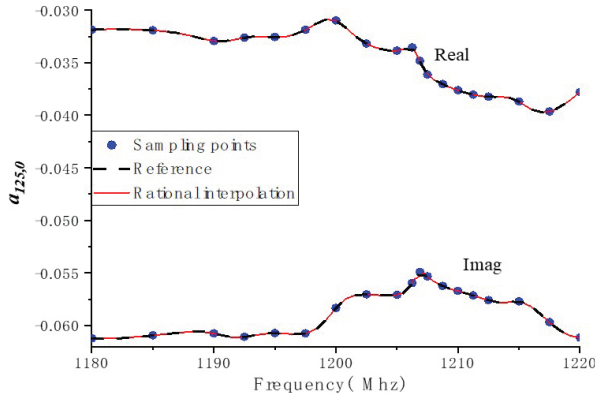


Fig. 14. Rational interpolation of coefficient $a_{125,0}$.

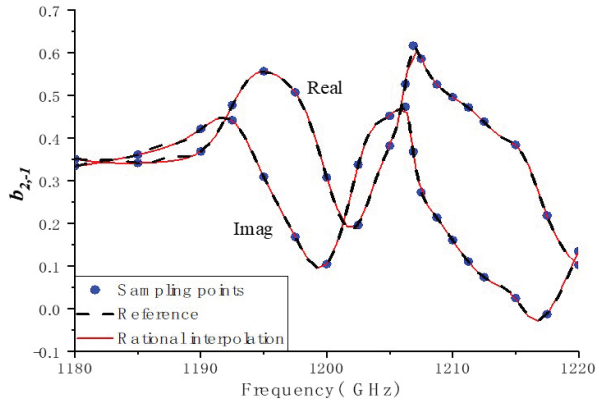


Fig. 15. Rational interpolation of coefficient $b_{2,-1}$.

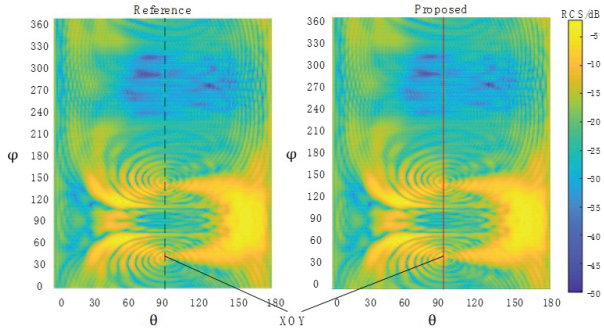


Fig. 16. Normalized bi-static RCS of UAV at 1216 MHz.

frequencies from 1180 MHz to 1220 MHz, and each frequency has 720 azimuthal angles and 360 elevational angles. The memory requirement of the complex vectors will be $40\times720\times360\times3\times16\approx498$ MB. The proposed algorithm needs only 19 frequencies over the whole band, and each frequency has 15761 non-zero spherical wave expansion coefficients. The memory requirement is $19\times15761\times16\approx4.78$ MB, and the compression ratio is about 103.

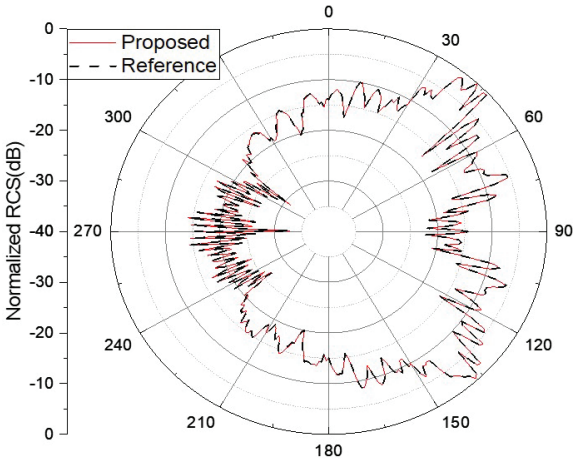


Fig. 17. Normalized bi-static RCS of UAV in XOY cut at 1216 MHz.

Finally, the efficiency of the proposed algorithm is compared to TDFT [10, 11] and SVD [12, 13] in Fig. 18. Figure 19 shows the spectrum obtained by TDFT, which converts a dense matrix into a sparse one by dropping the small elements in the spectrum matrix. If the relative error is set to be 0.1, the compression ratio of the proposed method is 64.13, while those of TDFT and SVD are 8.71 and 5.21, respectively. Thus, the proposed method is more efficient than the other two methods.

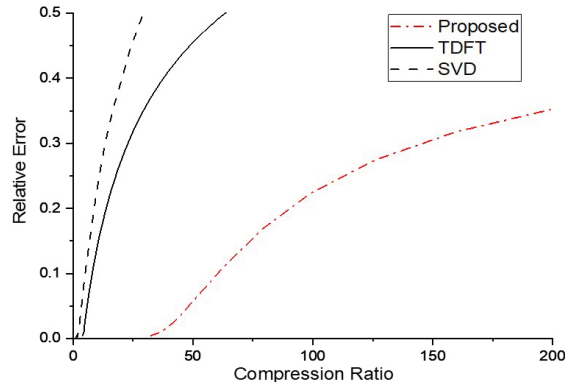


Fig. 18. Efficiency of three methods at 1216 MHz.

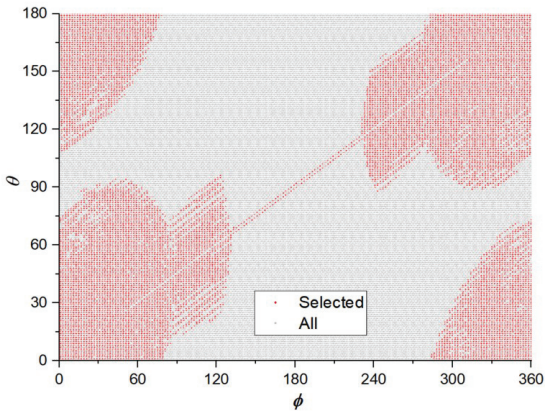


Fig. 19. Spectrum of the RCS obtained by TDFT.

IV. CONCLUSION

This paper gives a novel method to compress the field data of an antenna or a scatterer by SHT and rational interpolation. On the one hand, SHT converts the vectors of near field on a sphere into spherical wave expansion coefficients, which are low-pass sparse discrete spectra. On the other hand, rational interpolation fits these spectra over a frequency band with only a few sampling frequencies. As a result, the data of field are compressed dramatically, and we can readily restore the field data at an arbitrary frequency. This method can efficiently compress both far field and near field.

The proposed method will be improved in the future. First, the spherical wave expansion process will be replaced by FaVeST [24] or the spherical-multipole expansion [29]. Then, the proposed method will be revised to compressing the monostatic RCS, which is more useful in radar imaging and target recognition.

ACKNOWLEDGMENT

This work was supported in part by the Key Research and Development Plan of Shaanxi Province under Grant 2024GX-YBXM-232 and in part by the National Natural Science Foundation of China under Grant U2241203 and 62271365.

REFERENCES

[1] N. Y. Tseng and W. D. Burnside, "A very efficient RCS data compression and reconstruction technique, volume 4," NASA, Washington, D.C., Tech. Rep. NAS 1.26:191378-Vol-4, Nov. 1992.

[2] L. C. T. Chang, I. J. Gupta, W. D. Burnside, and C. L. T. Chang, "A data compression technique for scattered fields from complex targets," *IEEE Trans. Antennas Propagat.*, vol. 45, no. 8, pp. 1245-1251, Aug. 1997.

[3] I. J. Gupta and W. D. Burnside, "Electromagnetic scattered field evaluation and data compression using imaging techniques," NASA, Washington, D.C., Tech. Rep. NAS 1.26:207337, Sep. 1996.

[4] T. K. Sarkar and O. Pereira, "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials," *IEEE Trans. Antennas Propagat.*, vol. 37, no. 1, pp. 48-55, Feb. 1995.

[5] J. Tsao and B. D. Steinberg, "Reduction of sidelobe and speckle artifacts in microwave imaging: The CLEAN technique," *IEEE Trans. Antennas Propagat.*, vol. 36, no. 4, pp. 543-556, Apr. 1988.

[6] R. Bhalla and H. Ling, "Three-dimensional scattering center extraction using the shooting and bouncing ray technique," *IEEE Trans. Antennas Propagat*., vol. 44, no. 11, pp. 1445-1453, Nov. 1996.

[7] C. Wu, H. Zhao, and J. Hu, "Near field sampling compression based on matrix CUR decomposition," in *APS/URSI*, Singapore, pp. 1455-1456, 2021.

[8] H. Zhao, X. Li, Z. Chen, and J. Hu, "Skeletonization-scheme-based adaptive near field sampling for radio frequency source reconstruction," *IEEE Internet Things*, vol. 6, no. 6, pp. 10219-10228, Dec. 2019.

[9] H. Guo, Y. Liu, J. Hu, and E. Michielssen, "A butterfly-based direct integral-equation solver using hierarchical LU factorization for analyzing scattering from electrically large conducting objects," *IEEE Trans. Antennas Propagat*., vol. 65, no. 9, pp. 4742-4750, Sep. 2017.

[10] G. M. Binge and E. Micheli-Tzanakou, "Fourier compression-reconstruction technique (MRI application)," in *Images of the Twenty-First Century. Proceedings of the Annual International Engineering in EMBC*, Seattle, WA, pp. 524-525, 1989.

[11] W. X. Sheng, D. G. Fang, J. Zhuang, T. J. Liu, and Z. L. Yang, "Data compression in RCS modeling by using the threshold discrete Fourier transform method," *Chinese Journal of Electronics*, vol. 10, no. 4, pp. 557-559, Oct. 2001.

[12] S. J. Li, H. Pang, P. Y. Li, Y. N. Li, and Z. X. Liu, "Image compression based on SVD algorithm," in *CISAI*, Kunming, China, pp. 306-309, 2021.

[13] V. Cheepurupalli, S. Tubbs, K. Boykin, and N. Naheed, "Comparison of SVD and FFT in image compression," in *CSCI*, Las Vegas, Nevada, pp. 526-530, 2015.

[14] M. L. Don and G. R. Arce, "Antenna radiation pattern compressive sensing," in *MILCOM*, Los Angeles, CA, pp. 174-181, 2018.

[15] P. Debroux and B. Verdin, "Compressive sensing reconstruction of wideband antenna radiation characteristics," *Progress In Electromagnetics Research C*, vol. 73, pp. 1-8, 2017.

[16] M. Don, "Compressive antenna pattern measurement: A case study in practical compressive sensing," in *2022 IEEE AUTOTESTCON*, National Harbor, MD, pp. 1-9, 2022.

[17] W. Chen, C. Gao, Y. Cui, Y. Q. Jiang, and L. Yang, "A complex RCS calibration data interpolation method based on compressed sensing," in *PIERS - Fall,* Xiamen, China, pp. 1295-1299, 2019.

[18] H. Zhang, Y. Jiang, and X. Li, "Reconstruction of antenna radiation pattern based on compressed sensing," *Journal of Shanghai Jiaotong Univ. (Sci.),* vol. 25, pp. 790-794, July 2020.

[19] A. Massa, P. Rocca, and G. Oliveri, "Compressive sensing in electromagnetics: A review," *IEEE Trans. Antennas Propagat.*, vol. 57, no. 1, pp. 224-238, Feb. 2015.

[20] W. Dullaert and H. Rogier, "Novel compact model for the radiation pattern of UWB antennas using vector spherical and Slepian decomposition," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 2, pp. 287-299, 2010.

[21] B. Fuchs, L. le Coq, S. Rondineau, and M. D. Migliore, "Fast antenna far-field characterization via sparse spherical harmonic expansion," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 10, pp. 5503-5510, 2017.

[22] R. J. Allard and D. H. Werner, "The model-based parameter estimation of antenna radiation patterns using windowed interpolation and spherical harmonics," *IEEE Transactions on Antennas and Propagation*, vol. 51, no. 8, pp. 1891-1906, 2003.

[23] J. Gilewicz and Y. Kryakin, "Froissart doublets in Padé approximation in the case of polynomial noise," *Journal of Computational and Applied Mathematics*, vol. 153, no. 1-2, pp. 235-242, 2003.

[24] Z. Du, M. Kwon, D. Choi, and J. Koh, "Radiation pattern reconstruction techniques for antenna," *ICEIC*, pp. 143-146, 2008.

[25] J. E. Hansen, *Spherical Near-Field Antenna Measurements*. London: IET, 1988.

[26] Q. T. le Gia, M. Li, and Y. G. Wang, "Algorithm 1018: FaVeST-fast vector spherical harmonic transforms," *ACM Transactions on Mathematical Software*, vol. 47, no. 4, Dec. 2021.

[27] Y. Q. Xiao, S. Grivet-Talocia, P. Manfredi, and R. Khazaka, "A novel framework for parametric loewner matrix interpolation," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 9, no. 12, pp. 2404-2417, Dec. 2019.

[28] Y. Nakatsukasa, O. Sete, and L. N. Trefethen, "The AAA algorithm for rational approximation," *SIAM J. Sci. Comput*., vol 40, no. 3, pp. A1494-A1522, 2018.

[29] W. X. Sheng, D. G. Fang, J. Zhuang, T. Liu, and Z. L. Yang, "Comparative study on the data compression in RCS modeling," in *ISAPE*, Beijing, China, pp. 573-576, 2000.

[30] G. Giannetti and L. Klinkenbusch, "A numerical alternative for 3D addition theorems based on the bilinear form of the dyadic Green's function and the equivalence principle," *Advances in Radio Science*, vol. 22, pp. 9-15, 2024.

Haobo Yuan was born in Tianmen, Hubei, China, in 1980. He received the B.S., M.S., and Ph.D. degrees in electromagnetic fields and microwave technology from Xidian University, Xi'an, China, in 2003, 2006, and 2009, respectively. Since 2006, he has been with the School of Electronic Engineering, Xidian University, where he is an Associate Professor. He was a post-doctoral researcher at the Ohio State University in 2019. His current research interests include computational electromagnetics, antenna measurements, and electromagnetic compatibility.

Pu Yang was born in An'qing, Anhui, China, in 2000. He received the B.S. degree in electronic information engineering from Changchun University of Science and Technology, Changchun, China, in 2018. He is currently pursuing the M.S. degree with Xidian University. His research interests are numerical techniques in computational electromagnetics.

Nannan Wang received the B.S. degree in electronic information engineering from Nanjing University of Information Science and Technology, Nanjing, China, in 2023. She is currently pursuing the M.Eng. degree with Xidian University. Her research interest is antenna measurement techniques.

Yi Ren (Senior Member, IEEE) was born in Anhui, China, in 1982. He received the B.S. degree from Anhui University, Hefei, China, in 2004, and the Ph.D. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2009. In 2009, he joined Chongqing Jinmei Inc., Chongqing, China, as an Electromagnetic Compatibility Engineer. In 2011, he joined the Chongqing University of Posts and Telecommunications (CQUPT), Chongqing, where he was lately promoted as Full Professor. From 2014 to 2015, he was a Post-Doctoral Research Scholar with Duke University, Durham, NC, USA. Since August 2020, he has been a Professor with Xidian University, Xi'an, China. His research interests include computational electromagnetics and electromagnetic compatibility.

Shasha Li received the B.S. degree in electronic engineering from Northwestern Polytechnical University, Xi'an, China, in 2006. She received the Ph.D. degree in signal and information processing from Institute of Acoustics, Chinese Academy of Sciences (IACAS), Beijing, China, in 2012. Currently, she is an Engineer with Beijing Institute of Space Long March Vehicle, Beijing, China. Her current research interests include adaptive signal processing and array signal processing.

Design of 3-Bit Angle-insensitive RIS for 5G Communication Systems

T. Islam[1] and A. Eroglu[2]

[1]North Carolina A&T State University
Greensboro, NC 27411, USA
tislam5@aggies.ncat.edu

[2]SUNY Polytechnic Institute
Utica, NY 13502, USA
eroglua@sunypoly.edu

Abstract – This paper introduces a 3-bit reconfigurable intelligent surface (RIS) design characterized by its unique angle-insensitive properties for 5G communication systems. The proposed configuration provides eight distinct phase states enabled by the states of two varactors with an applied bias voltage. The design of the unit cell with double centric square split ring resonators and the formation of the RIS with a 5x5 array have been presented. A detailed analysis of the RIS performance has been conducted using the CST 3D electromagnetic simulator to study the reflection amplitude and phase responses. It is demonstrated that the results show a phase range of up to 315 degrees, along with eight distinct states exhibiting a stable interval of 45 degrees. This effectively covers incidence angles ranging from 0 to 60 degrees.

Index Terms – 3-bit RIS, 5G, angular insensitivity, quantization, reconfigurable intelligent surface, split ring resonator, SRR, varactor.

I. INTRODUCTION

The wireless communication landscape is currently undergoing a transformative phase with the emergence of the reconfigurable intelligent surface (RIS) [1]. These innovative platforms are gaining widespread recognition for their ability to effectively manipulate signal propagation environments, thereby significantly enhancing wireless network performance. The traditional challenges associated with uncontrolled interaction of radio waves environmental elements, which often lead to signal quality deterioration, are being addressed through the deployment of RIS [2]. These surfaces play a pivotal role in optimizing wave reflections, refractions, and scattering, which in turn facilitates the reduction of multipath fading while maintaining low hardware costs and energy efficiency [3–4].

At the core of RIS technology is the metasurface, a crucial component renowned for its remarkable capability to manipulate electromagnetic (EM) waves through meticulously designed and strategically arranged meta-atoms [5]. The concept of digital coding metasurfaces has revolutionized the design process, allowing for real-time programmability and reconfigurability. The integration of tunable components like PIN diodes, varactor diodes, and field-programmable gate arrays (FPGA) has further enhanced the functional diversity of metasurfaces, paving the way for high-performance multifunctional metasurfaces in RIS-assisted wireless communication [6].

This paper introduces a 3-bit programmable RIS to address the challenges arising from angular sensitivity and phase control in RIS. Such sensitivity may lead to issues that can result in the failure of RIS-assisted wireless communication networks such as the ones that rely on the reciprocity of wireless channels [7–8]. The proposed design in this paper features a unit cell which has double centric square split ring resonators (SRRs) with two varactors for phase control, aiming to enhance angular insensitivity [9–10]. The unit cell is then used to form RIS with 5×5 array configuration. The results demonstrate a phase range of up to 315 degrees, with eight distinct states exhibiting a stable interval of 45 degrees, effectively covering incidence angles from 0 to 60 degrees [11–12].

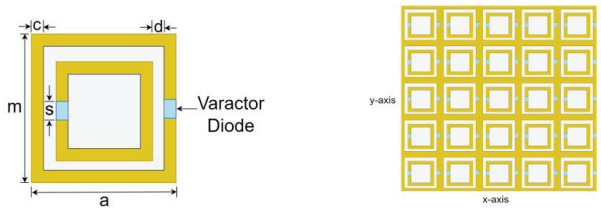
II. DESIGN OF SRR UNIT CELL

The square SRR unit cell in the design is implemented on a Teflon-based substrate, as depicted in Fig. 1 (a). Each SRR unit cell comprises a square loop with a side length (a) of 5 mm, a split width (d) of 0.5 mm, and a metal width (c) of 0.2 mm, m represents the side length of the SRR unit cell and is equal to 5 mm. The spacing (s) between adjacent SRRs is maintained at 0.5 mm, optimizing the coupling effect and resonance characteristics. A detailed analysis has been conducted and verification using the 3D EM simulator CST. The varactor diode model used is the MAVR-00020-141100 from

Submitted On: September 2, 2024
Accepted On: February 18, 2025

https://doi.org/10.13052/2025.ACES.J.400406

MACOM. The capacitances of this varactor diode range from 0.75 pF to 2.60 pF.



(a) (b)

Fig. 1. Layout of SRR unit cell with varactor diodes: (a) single element unit cell and (b) 5x5 array with final unit cell.

III. SIMULATION RESULTS AND DISCUSSION

Figure 2 illustrates the normalized reflection coefficient amplitude across a frequency range from 20 GHz to 30 GHz for various varactor capacitance values. Each curve represents the response for a specific capacitance value and its corresponding phase shift, demonstrating the shift in the resonant frequency of the element with changes in capacitance. The peak amplitude values in Fig. 2 indicate the resonant points where maximum reflection occurs. These characteristics demonstrate the ability to control the phase of reflected signals by adjusting the varactor capacitance, a key characteristic for beam-steering applications.

Figure 3 displays the linear phase shift progression of a SRR unit cell at 26 GHz, with the reflection phase angle increasing in 45° increments across eight configuration states, ranging from 0° degrees at state 1 to 315°

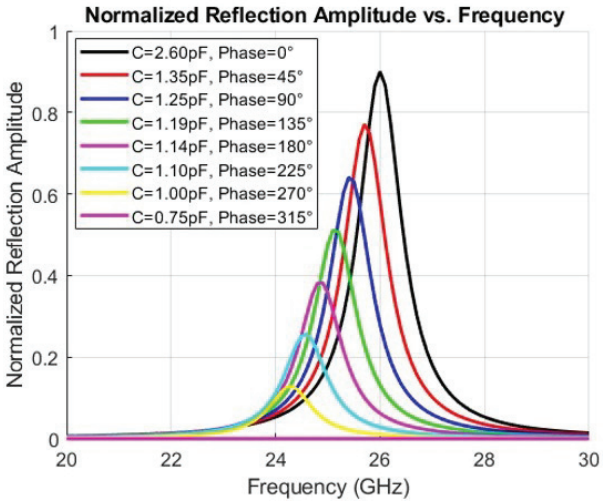


Fig. 2. Frequency response of a varactor-tuned RIS.

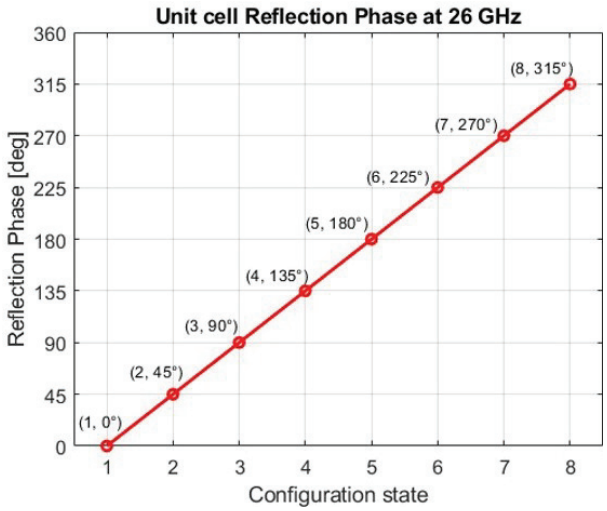


Fig. 3. Progressive phase shift characteristics of SRR unit cell at 26 GHz.

degrees at state 8. In Fig. 4, the variation of the normalized scattering pattern for different incident angles over an 180° degree range is given, with each curve representing the scattering intensity at a specific incident angle.

RIS angular reciprocity is investigated to identify the angular sensitivity of the proposed RIS structure. Sets of simulations have been conducted for the predetermined incident angles and corresponding reflected angles have been measured and illustrated in Table 1.

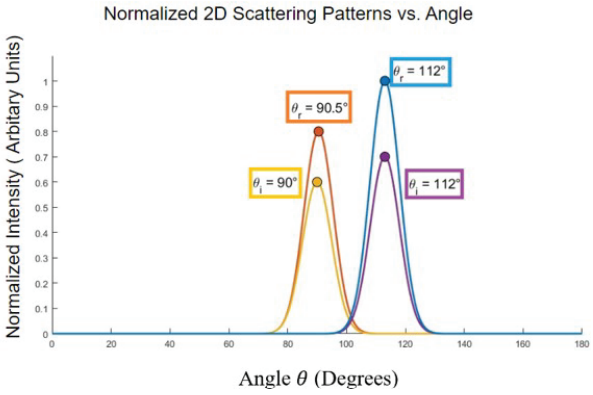


Fig. 4. Normalized 2D scattering pattern vs. angle.

Table 1: Angular sensitivity

Simulation I	Incident Angle - Θ_i	90°	112.5°
	Reflected Angle - Θ_r	112.5°	90.5°

The normalized 2D scattering parameter responses for each incident angle are illustrated in Fig. 4. The simulation results affirm that plane waves are reflected back

toward the incidence direction on the RIS, thereby validating the concept of angular reciprocity.

The DC biasing circuit for the varactor is given in Fig. 5. Figure 6 gives the capacitance of the varactor versus applied reverse voltage. The complete system with the varactor bias circuit is integrated and simulated as shown in Fig. 7. The return loss for the unit cell when the applied reverse voltage is between 0.5 V and 5.6 V is plotted in Fig. 8. Figure 8 illustrates how the return loss varies with the frequency for each bias setting, indicating the varactor's performance and tenability within this range.

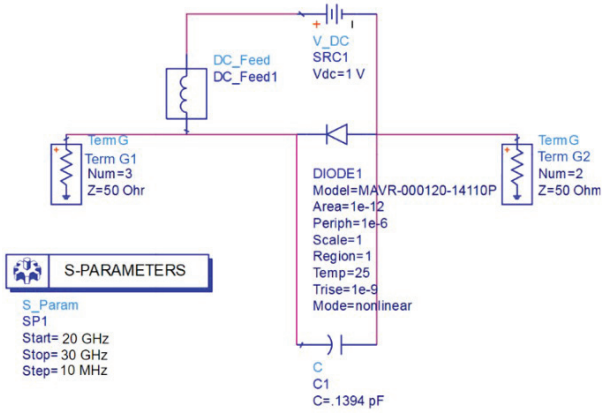


Fig. 5. DC biasing circuit for the presented unit cell reconfigurable diode.

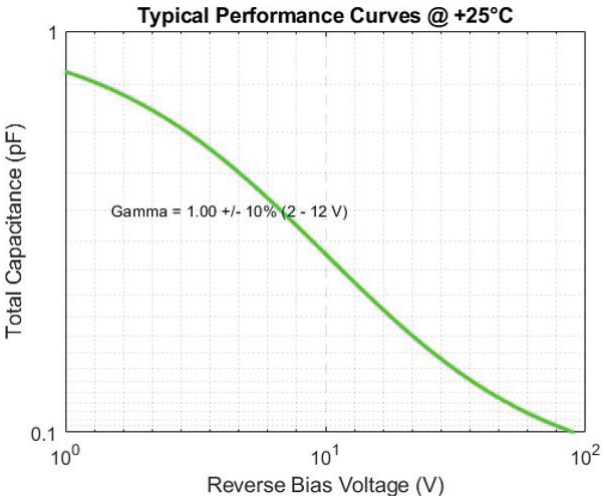


Fig. 6. Capacitance of the varactor versus reverse bias voltage.

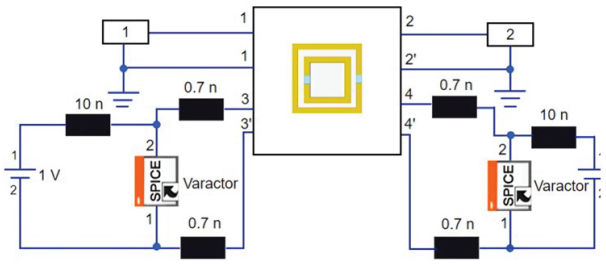


Fig. 7. Integrating the SRR unit cell with the DC circuit and SPICE representation for varactor diodes.

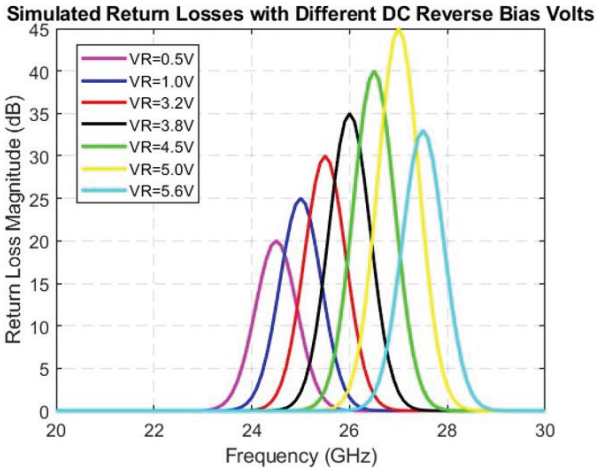


Fig. 8. Simulated return losses with different reverse bias voltages.

Figure 9 illustrates the insertion loss for the applied reverse voltages. Figure 10 gives the group delay for the unit cell when reverse bias voltages vary. The phase characteristics of the insertion loss is given in Fig. 11. Figure 12 shows two color-coded plots representing phase distributions for a RIS. The top plot is labeled "Continuous phase distribution" and shows a gradient of colors, indicating a smooth transition of phase values across the surface. The bottom plot is labeled "2-

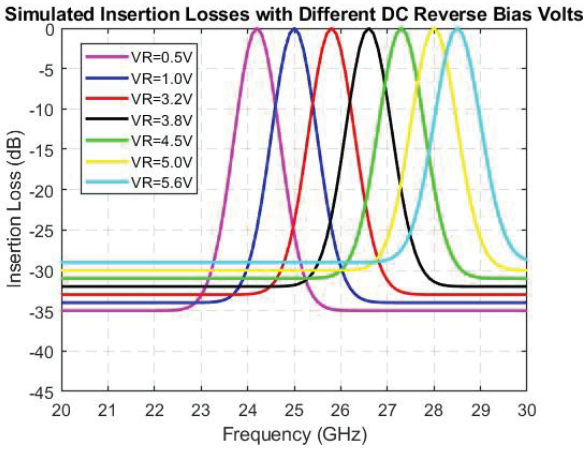


Fig. 9. Insertion losses with different reverse bias voltages.

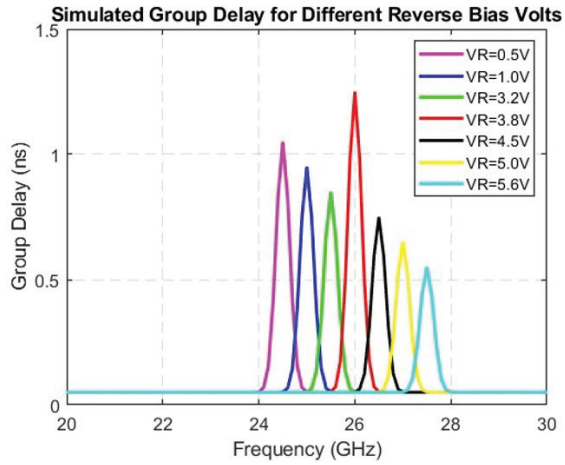


Fig. 10. Simulated group delay for different reverse bias voltages.

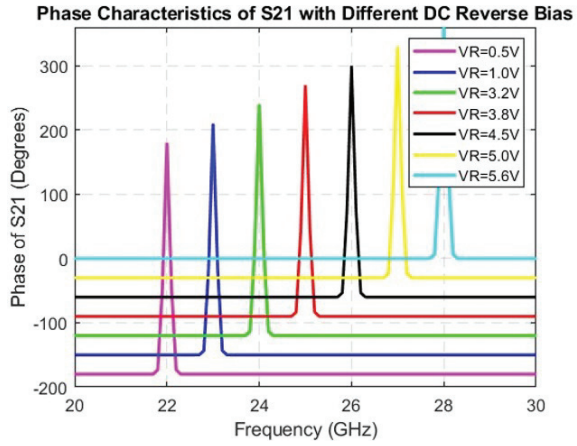


Fig. 11. Simulated phase characteristics of S21 with different DC reverse bias voltages.

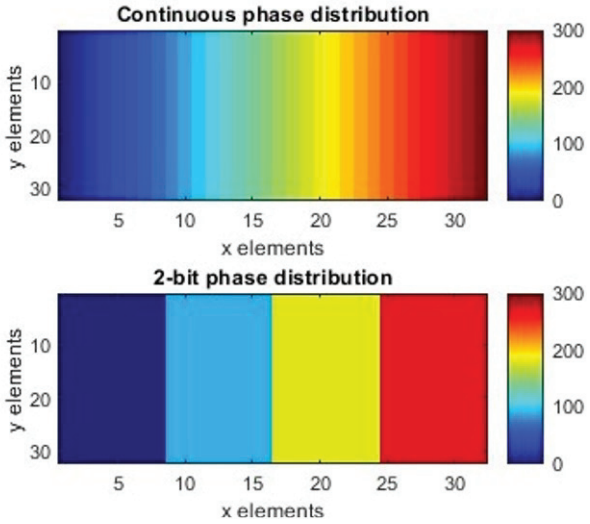


Fig. 12. Comparative displays of continuous and 2-bit phase distribution.

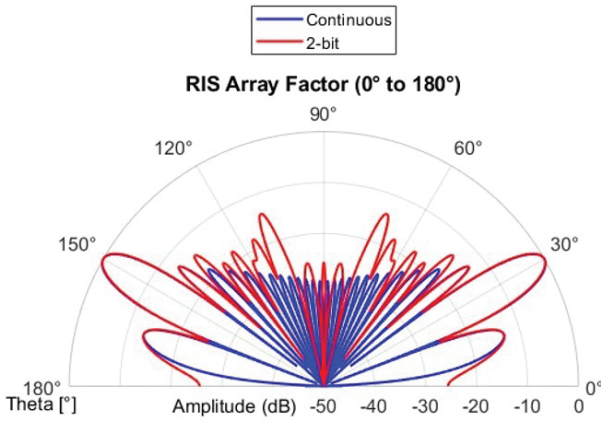


Fig. 13. Comparison of continuous and 2-bit phase control for RIS beam steering at 26 GHz.

bit phase distribution" and displays four distinct colors, representing discrete phase states achievable with 2-bit quantization.

Figure 13 illustrates the RIS array factor for beam steering, comparing continuous and 2-bit phase control methods. It displays the amplitude in decibels (dB) against the angle theta in degrees. The continuous phase control (blue line) results in a smoother, more precise, steering profile, while the 2-bit phase control (red line) shows a more discretized pattern with higher side lobes, indicating less precise beam control.

Figure 14 displays four plots that compare the performance of an RIS array factor at 26 GHz with the different quantization levels such as 1-bit, 2-bit, and 3-bit. The top left plot indicates the mean quantization phase error which measures the average phase error due to quantization at various reflected angles. The top right plot represents the quantization loss. This plot indicates the power loss associated with phase quantization. The bottom left plot represents the beam pointing error which shows the deviation of the actual beam direction from the intended direction due to quantization. The bottom right plot shows the side lobe level. This plot represents the relative power level of side lobes compared to the main lobe, a critical factor in beamforming performance. Figure 15 contains two graphs depicting the performance of an RIS array at 26 GHz based on the different unit cell sizes (0.1λ, 0.25λ, and 0.5λ where λ is the wavelength). The left graph shows the pointing error in degrees as a function of the reflected angle. It demonstrates how the beam's actual direction deviates from the targeted direction for each unit cell size. The right graph illustrates the level of the side lobes in dB relative to the main lobe, also as a function of the reflected angle, indicating how

unit cell size impacts the prominence of the side lobes during beamforming.

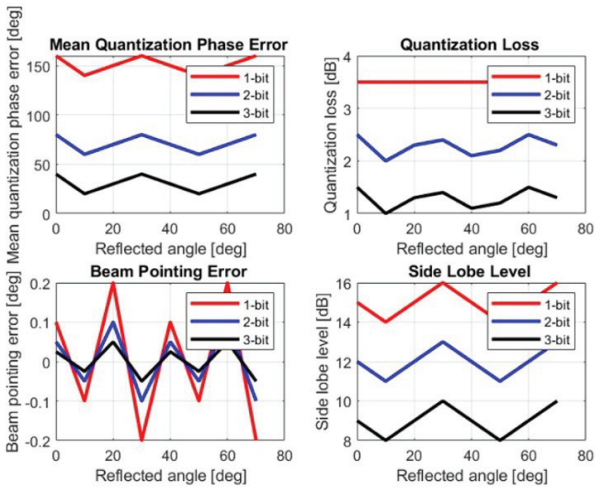


Fig. 14. Performance metrics of RIS array factor with phase quantization at 26 GHz.

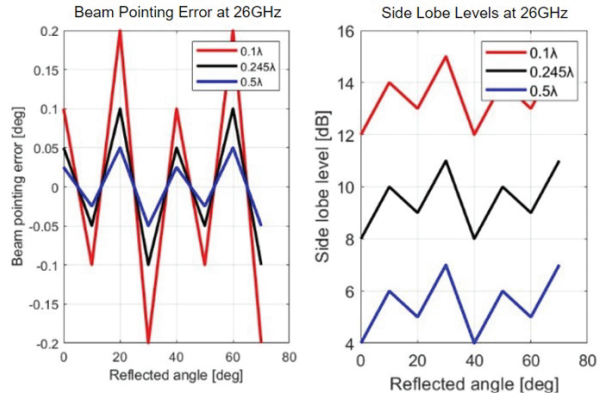


Fig. 15. Impact of unit cell size on RIS beam pointing accuracy and side lobe levels at 26 GHz.

IV. CONCLUSION

In this paper, we have designed and simulated a 3-bit reconfigurable intelligent surface (RIS). The proposed design incorporates a unit cell with angular reciprocity, aiming to enhance uplink and downlink misalignment. This improvement addresses the potential risk of failure in RIS-assisted wireless communication networks caused by misalignment issues. It is shown that the RIS presented in this paper achieves a comprehensive phase shift range from 0° to 315°, enabling eight distinct digital states with stable 45° intervals. This has been accomplished by incorporating two varactors and enabling eight distinct phase states by changing the applied voltage. The outcomes of this research can be used for communication systems that rely on precise angular sensitivity and critical phase control.

REFERENCES

[1] Q. Ma, Y. Liu, Z. Zhang, X. Wang, J. Wang, and H. Li, "Smart metasurface with self-adaptively reprogrammable functions," *Light: Science & Applications*, vol. 8, no. 1, Oct. 2019.

[2] L. Zhang, X. Q. Chen, S. Liu, Q. Zhang, J. Zhao, and T. J. Cui, "Dynamically realizing arbitrary multi-bit programmable phases using a 2-bit time-domain coding metasurface," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 4, pp. 2984-2992, Apr. 2020.

[3] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. D. Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 421-439, Jan. 2021.

[4] A. Eroglu, *RF/Microwave Engineering and Applications in Energy Systems*. Hoboken, NJ: Wiley, 2022.

[5] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting-surface: Practical phase shift model and beamforming optimization," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5849-5863, Sep. 2020.

[6] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. D. Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450-2525, Nov. 2020.

[7] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78729-78757, 2019.

[8] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, and Z. Chen, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45913-45923, 2020.

[9] S. R. Biswas, C. E. Gutierrez, A. Nemilentsau, I. H. Lee, S. H. Oh, P. Avouris, and T. Low, "Tunable graphene metasurface reflectarray for cloaking, illusion, and focusing," *Phys. Rev. Appl.*, vol. 9, no. 034021, 2018.

[10] A. Araghi, M. Khalily, M. Safaei, A. Bagheri, V. Singh, F. Wang, and R. Tafazolli, "Reconfigurable

intelligent surface (RIS) in the sub-6 GHz band: Design, implementation, and real-world demonstration," *IEEE Access*, vol. 10, pp. 2646-2655, 2022.

[11] A. Pitilakis, O. Tsilipakos, F. Liu, K. M. Kossifos, A. C. Tasolamprou, D. H. Kwon, M. S. Mirmoosa, and C. M. Soukoulis, "A multi-functional reconfigurable metasurface: Electromagnetic design accounting for fabrication aspects," *IEEE Trans. Antennas Propag.*, vol. 69, no. 3, pp. 1440-1454, 2021.

[12] O. M. Bucci, G. Mazzarella, and G. Panariello, "Reconfigurable arrays by phase-only control," *IEEE Trans. Antennas Propag*., vol. 39, no. 7, pp. 919-925, 1991.

Tamanna Islam is currently pursuing a Ph.D. in Electrical Engineering at North Carolina A&T State University. She received her M.S. in Electrical Engineering from the same institution in 2023 and her B.S. in Electrical Engineering from the Military Institute of Science and Technology in Bangladesh in 2019. Her professional experience includes work on reconfigurable intelligent surfaces (RIS), phased array antennas, THz sensors, and AI-integrated wireless communication systems. Her research interests include RIS design, THz device development, AI-based beamforming, and adaptive signal processing for 5G and 6G communication systems.

Abdullah Eroglu received his Ph.D. in Electrical Engineering from the Electrical Engineering and Computer Science Department at Syracuse University, Syracuse, NY, USA, in 2004. He was a Senior RF Design Engineer at MKS Instruments from 2000 to 2008, where he was involved in the design of RF systems. He is currently the Dean of the College of Engineering at SUNY Polytechnic Institute and an Emeritus Professor at Purdue University. He was a Faculty Fellow with the Fusion Energy Division at Oak Ridge National Laboratory, Oak Ridge, TN, USA, in 2009. He has published over 170 journal and conference papers and holds five patents. He is the author of six books and co-editor of one book. His current research interests include RF/microwave/THz circuit design and applications, anisotropic, gyrotropic, and metamaterials, RF amplifiers and topologies, antennas, RF metrology, and the wave propagation and radiation characteristics of materials.

Beam-reconfigurable Antenna Based on Planar Inductor with Mn-Zn Ferrite

Meng Yin, Cheng Ju, Wei-Hua Zong, and Shandong Li

College of Electronic Information
Qingdao University, Qingdao 266071, China
2018206231@qdu.edu.cn, jucheng719@qdu.edu.cn, zongweihua@qdu.edu.cn, lishd@qdu.edu.cn

***Abstract* –** This paper presents a beam-reconfigurable antenna design adopting distributed inductors, Mn-Zn ferrite, and static magnetic fields. The proposed antenna consists of one driven patch, two parasitic patches, and a full ground plane. Each parasitic patch is loaded with a distributed inductor with positive inductance. The patch antenna has a symmetric configuration and a broadside pattern. A Mn-Zn ferrite slab is added to one inductor to reduce its self-resonant frequency and change its inductance from positive to negative which results in unsymmetric field distributions and a tilted radiation beam. A static magnetic field is applied to the ferrite material further to adjust the tilted angle of the radiation beam. The proposed antenna works at five modes with reconfigurable beams of θ =0° (ϕ =0°), 15° (ϕ =90°, 270°) and 28° (ϕ =90°, 270°).

***Index Terms* –** Magnet, Mn-Zn ferrite, negative inductances, reconfigurable beams.

I. INTRODUCTION

Beam-reconfigurable antennas provide interference-free, power-saving, and highly secured end-to-end communication making them attractive in 5G and satellite communication systems [1–4]. Extensive research has been conducted on beam reconfigurable antenna design [3–13], in the design of which PIN diodes [3–9], varactors [10, 11], and phase shifters [12, 13] are adopted to control feeding networks, connections among metallic portions, or phase shifting between antenna elements. Table 1 compares the published beam reconfigurable antennas. A common characteristic of these designs is the need to introduce a DC power supply for control. This requires the design of DC bias circuits, which involves a large number of lumped components, increases design complexity, occupies space on the circuit board, and potentially introducing losses that lead to gain reduction during the soldering of these lumped components.

To address these issues, this article presents a novel technique of beam reconfiguration using tunable inductors. The inductors are loaded with magnetic material and controlled by a magnetic field to tune both the inductance and the antenna beam. Compared to traditional beam reconfigurable antennas, the proposed design does not require a DC bias circuit, and the fabrication only involves grounding the inductor, simplifying the manufacturing process, and reducing design complexity. Magnetic materials and magnetic fields have been adopted in inductors to improve [14–19] and tune [20–23] inductance. To the best of the authors' knowledge, this article presents the first design of a beam-reconfigurable antenna by using magnetic field control.

Table 1: Comparison of the published beam reconfigurable antennas

Ref.	Freq. (GHz)	Tunable and Lumped Components	Beam States	Gain (dB)
[3]	3.5	4 PIN, 4 L, 4 R, 4 P	8	4.9
[4]	3.8	8 PIN, 16 L, 32 C, 8 R, 8 P	3	3.8
[5]	3.6	8 PIN, 8 L, 8 C, 8 R, 8 P	5	11.85
[6]	3.7	12 PIN, 4 L, 12 P	12	4.61
[7]	5.3	8 PIN, 9 L, 12 C, 1 R, 8 P	4	7.04
[10]	3	2 Var, 2 C, 2 P	5	N.A.
[11]	5.8	3 Var, 3 L, 3 C, 3 R, 3 P	3	6.5
[12]	11.75	Phase shifter	6	20
[13]	4.8	28 PIN, 12 C, 28 P	9	7.8[§]
This Work	2.4	0	5	6.02

PIN is PIN diode, L is inductor, C is capacitor, R is resistor, P is power, Var is varactor, [§]unit is dBic

The proposed antenna is a patch antenna with two parasitic patches. Two tunable inductors are loaded in the two parasitic patches. A Mn-Zn ferrite slab and magnet are adopted to tune each inductor's inductance. The proposed antenna beam works at five states with main beam points to θ =28° (ϕ =270°), 15° (ϕ =270°), 0° (ϕ =0°), 15° (ϕ =90°), and 28° (ϕ =90°), respectively.

II. ANTENNA DESIGN

Figure 1 shows the configuration of the proposed antenna. The antenna consists of three rectangular

Submitted On: October 30, 2024
Accepted On: March 31, 2025

https://doi.org/10.13052/2025.ACES.J.400407

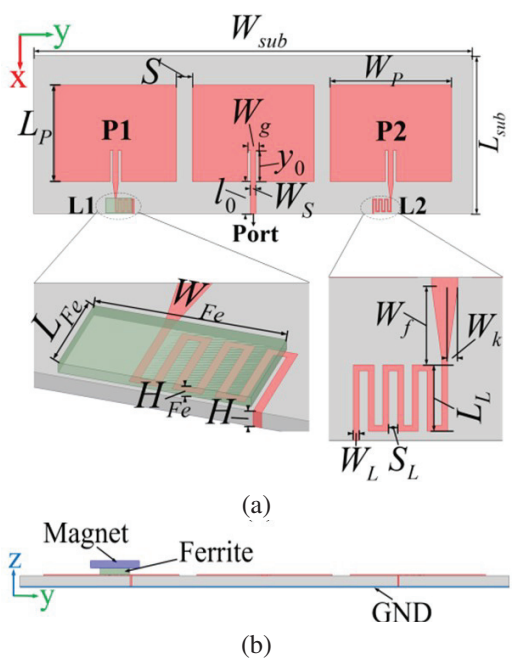


Fig. 1. Configuration of the proposed antenna: (a) top view (without magnet) and (b) side view (with magnet).

patches and a full ground (GND) printed on a 0.8 mm thick Rogers RO4003 substrate with ε_r =3.55 and tanδ =0.0027. The three patches are identical in size and shape. The center patch is directly fed through a microstrip line, while the other two patches (P1 and P2) are fed by coupling, with each loaded with a meander-shaped inductor shorted to the ground plane. The antenna configuration derives from the one in [10] by replacing the varactors with tunable inductors. The patch antenna is symmetric along the y-axis, resulting in a broadside radiation pattern. A piece of Mn-Zn ferrite slab, with a thickness of 0.5 mm, is mounted on top of one inductor to realize negative inductance, while the other inductor, which has no ferrite slab, exhibits positive inductance. The discrepancy in inductance values between the two inductors results in a tilted beam. Furthermore, applying a magnetic field with a magnet results in a larger scanning beam. The proposed antenna operates at a resonant frequency of 2.4 GHz. The antenna dimensions are listed in Table 2.

Table 3 shows the radiation patterns of the proposed antenna. At Mode 1, both inductors are without a ferrite slab or magnet, resulting in equal inductance values and broadside patterns. At Mode 2, the inductance on the right side is loaded with a ferrite slab, resulting in a beam tilted θ =15° and ϕ =90°. At Mode 3, the antenna beam tilts to the opposite direction of Mode 2 (θ =15° and ϕ =270°) with the ferrite slab loaded on the left inductor. At Mode 4/5, the ferrite slab on the right/left inductor is subjected to a magnetic field from a magnet, and a larger tilted beam is obtained at θ =28° and ϕ =90°, 270°.

A. Reconfigurable beam design with lumped inductor

Two lumped inductors are used to replace the two meandered line inductors (L1 connected to P1 and L2 connected to P2) to evaluate the inductance values required to obtain beam reconfiguration. Antenna performance was simulated using the Ansys high-frequency structure simulator (HFSS) software.

Table 4 presents 14 states with different beams and Fig. 2 illustrates the corresponding radiation patterns. As shown in Table 4 and Fig. 2, at State 1, L1=L2=10 nH, the antenna beam points to the +z direction (θ =0° and ϕ =0°). At State 2, L1 and L2 are both positive with a great difference in value (9999 nH), and a small tilted

Table 2: Dimensions of the proposed antenna

Parameter	W_{sub}	L_{sub}	W_P	L_P	W_S	W_g	y_0	l_0	s
Value (mm)	150	53	41.5	32.8	1.8	3.7	10.6	11	5.5
Parameter	H	W_{Fe}	L_{Fe}	H_{Fe}	W_L	L_L	S_L	W_f	W_k
Value (mm)	0.8	5	9	0.5	0.4	3	0.5	5.5	1.9

Table 3: Radiation patterns of the proposed antenna

Mode	**Left Inductor**	**Right Inductor**	**Measurement**		**Simulation**	
			Main Beam (θ,φ)	**Gain (dB)**	**Main Beam (θ,φ)**	**Gain (dB)**
1	-	-	(0°, 0°)	6.82	(0°, 0°)	6.84
2	-	With ferrite	(15°, 90°)	6.22	(15°, 90°)	6.58
3	With ferrite	-	(15°, 270°)	6.22	(15°, 270°)	6.58
4	-	With ferrite and magnet	(28°, 90°)	6.07	-	-
5	With ferrite and magnet	-	(28°, 270°)	6.07	-	-

angle of 1° is obtained in the elevation plane. This indicates that loading inductors with different values in the two parasitic patches introduces a beam tilted to the side of the smaller inductor. However, a minimal tilted angle is obtained when both inductors are positive. In the cases of States 3-6, one of the two patches is directly shorted to the ground plane without loading an inductor, while the other patch is loaded with an inductor. A negative inductor (-10 nH) achieves a larger tilted angle compared to a positive inductor (100 nH), indicating that negative inductors are more effective in achieving larger tilted beams than positive ones. The comparison of States 7-10 indicates that when the inductance values of the two inductors have opposite signs, a smaller absolute value of the negative inductance leads to a larger tilted angle in the radiation pattern. In States 9-10, a significant tilted angle was achieved, θ =32°, where one inductor is a positive inductance of 260 nH and the other is a negative inductance of -3 nH. Comparing State 10 with States 11-14, it can be seen that the titled beam angle increases with the increase of the positive inductor's inductance value or the negative inductor's inductance absolute value. It seems that the beam angle reaches a stable value at State 9-10, since the beam angle only increases by 1° when the positive inductor increases from 260 to 460 nH or the negative inductor decreases from -3 to -1 nH. In section IIB, a distributed inductor with an inductance around 260 nH will be designed using meandered lines, and the inductance is tuned to a negative value around -3 nH by loading a ferrite slab.

Table 4: States of proposed antenna with lumped inductors

State	L1 (nH)	L2 (nH)	Beam direction (θ,φ)
1	10	10	(0°, 0°)
2	1	10000	(1°, 270°)
3	0	100	(4°, 270°)
4	100	0	(4°, 90°)
5	-10	0	(10°, 270°)
6	0	-10	(10°, 90°)
7	-20	260	(13°, 270°)
8	260	-20	(13°, 90°)
9	-3	260	(32°, 270°)
10	260	-3	(32°, 90°)
11	160	-3	(30°, 90°)
12	360	-3	(32°, 90°)
13	460	-3	(33°, 90°)
14	260	-1	(33°, 90°)

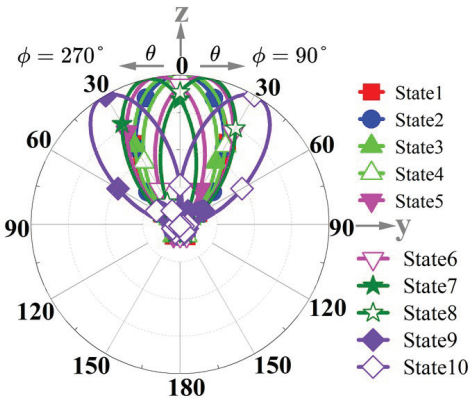


Fig. 2. Normalized radiation patterns for different states of yoz.

B. Design of distributed inductors

Based on the discussion above, different tilted beams are obtained by loading different inductors in P1 and P2. Meandered lines are utilized to design a positive inductor, avoiding the need for via holes or metallic bridges that would be required for a spiral-shaped inductor [17]. In the proposed inductor, shown in Fig. 1, one terminal is directly connected to the parasitic patch, while the other terminal, located at the edge of the board, connects to the ground plane via a soldered copper wire. Figure 3 shows the inductor's characteristics varying with dimensions. As shown, the inductor reaches its maximum positive value at its self-resonant frequency and decreases to a negative value dramatically with the increase in frequency. Furthermore, the inductor's resonant frequency decreases with the increase of the line's length of L_L, width of W_L, and spacing of S_L. The inductor resonates at the working frequency of 2.4 GHz when L_L =3 mm, W_L =0.4 mm, S_L =0.5 mm, with its peak inductance of 260 nH which meets the requirement of positive inductor's inductance value at States 9-10 in Table 4. A negative inductor with an inductance around -3 nH is required based on the meandered line-shaped configuration working at States 9-10.

Based on the inductors designed above, this paper presents an approach of using magnetic material to change the inductance from positive to negative by reducing the inductor's self-resonant frequency [17, 24]. Mn-Zn ferrite is chosen to be loaded onto one inductor as shown in Fig. 1 (a), due to its high magnetic permeability and sensitive response to magnetic fields. Figure 4 shows the inductor's inductance with/without the ferrite slab with W_L =0.4 mm and S_L =0.5 mm. When L_L =3 mm, with a loaded ferrite slab, the self-resonant frequency of inductors decreases from 2.4 GHz to 2.1 GHz, while its inductance changes from 260 nH to -20 nH. When L_L =2 mm, both resonate frequencies are higher than 2.4 GHz, resulting in two positive inductances of the cases without/with ferrite slab. When L_L =4 mm, both resonate frequencies are lower than 2.4 GHz, resulting in two negative inductances of the cases without/with ferrite slab.

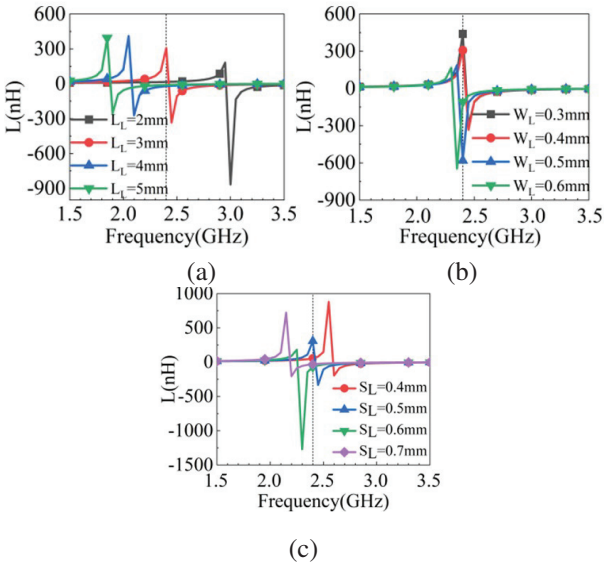


Fig. 3. Impact of parameter dimensions on inductance: (a) L_L, (b) W_L, and (c) S_L.

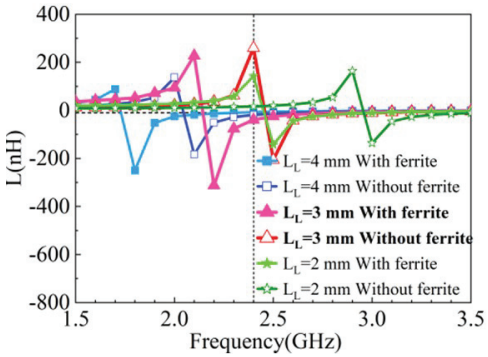


Fig. 4. Changes of inductance by adding ferrite slab.

In conclusion, an inductor resonating at its working frequency (2.4 GHz) is an ideal choice to obtain negative inductance by adding a ferrite slab.

Figure 5 gives a simulated magnetic field distribution on the inductor with/without ferrite slab. As shown, the magnetic fields increase significantly with the loaded ferrite slab, which leads to a decrease in self-resonant frequency and an increase in inductance at frequencies below resonant frequencies, which agrees with the trend shown in Fig. 5.

C. Reconfigurable beam design with distributed inductors

By adopting the two distributed inductors with meandered line shapes and loading with a ferrite slab (as shown in Fig. 1), beam reconfiguration is obtained with three modes given in Table 2. At Mode 1, both inductors are unloaded with a ferrite slab, resulting in the same

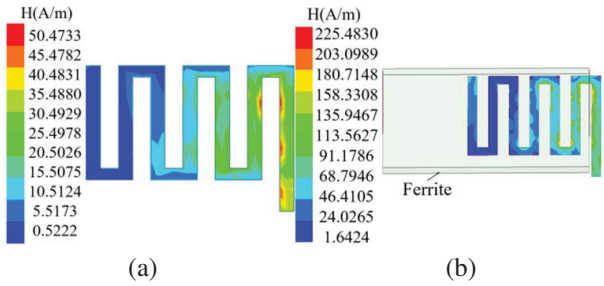


Fig. 5. Simulation of H-field distribution of the proposed inductor: (a) without ferrite and (b) with ferrite.

inductance and broadside pattern. At Mode 2, the right inductor has negative inductance by loading with a ferrite slab, and the left one has positive inductance without ferrite loading, resulting in the main beam being tilted 15° to the negative inductor. At Mode 3, the antenna beam is tilted to the left side when a ferrite slab is loaded on the left inductor.

S_{11} plots of the proposed antenna in the three modes are given in Fig. 6. The antenna resonates at 2.4 GHz with an overlapping bandwidth of 2.39-2.42 GHz in the three modes. Figure 7 presents the proposed antenna's simulated electric field distributions and radiation patterns at different modes. At Mode 1, the electric field distributions and radiation patterns show symmetry about the antenna's middle axial due to its symmetric configuration. At Mode 2/3, the electric field is more concentrated around the patch loaded with a ferrite slab, resulting in the antenna beam being tilted to the same side.

Table 5 illustrates the impact of varying the ferrite slab's dimensions on the antenna's radiation pattern. Specifically, as the thickness (H_{Fe}) of the ferrite slab increases, the tilted beam angle increases while the gain decreases. To achieve an optimal balance between the beam angle and gain, an H_{Fe} of 0.5 mm was selected.

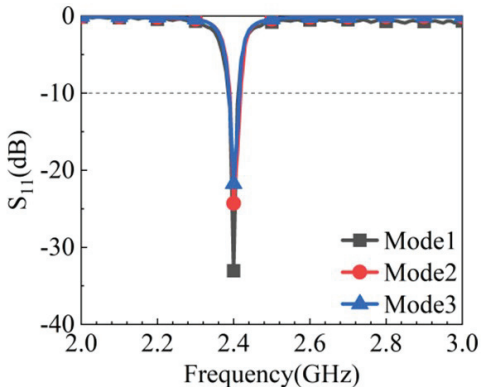


Fig. 6. S_{11} plots of the proposed antenna.

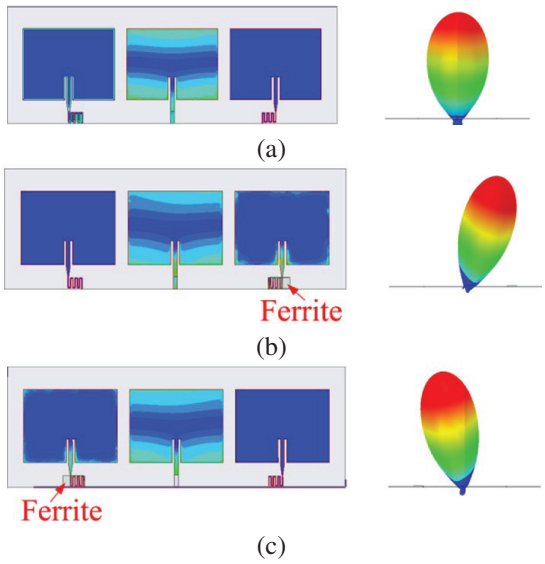


Fig. 7. E-field distribution and radiation pattern: (a) Mode 1, (b) Mode 2, and (c) Mode 3.

Additionally, the maximum beam angle and gain were observed when W_{Fe} =5 mm and L_{Fe} =9 mm.

Table 5: Ferrite slab's size influence on antenna pattern

Parameter	Value (mm)	Beam (θ,φ)	Gain (dB)
$\mathbf{H}_{Fe}$	0.45	(13°, 90°)	5.0
	0.5	**(18°, 90°)**	**4.9**
	0.55	(20°, 90°)	4.6
$\mathbf{W}_{Fe}$	4.5	(13°, 90°)	4.8
	5	**(18°, 90°)**	**4.9**
	5.5	(15°, 90°)	4.6
$\mathbf{L}_{Fe}$	8.5	(15°, 90°)	4.7
	9	**(18°, 90°)**	**4.9**
	9.5	(16°, 90°)	4.6

Based on the discussion of section IIA, an increase in the value of negative inductance will lead to a larger tilted beam angle. A permanent magnet was adopted to enforce static magnetic fields on the Mn-Zn ferrite. Magnetic fields distributions in the ferrite slab with/without the magnet have been simulated by using ANSYS Maxwell and are given in Fig. 8. As shown, by applying static magnetic fields, the magnetic induction intensity increases more than two times. The effect of the magnet on the inductor's inductance and the tilted beam cannot be simulated by ANSYS Maxwell or HFSS directly due to the limitations of each software. Maxwell can only model the impact of static magnetic fields on the magnetic properties but cannot account for the dynamic behavior of the inductor or antenna at higher frequencies. HFSS can accurately model the inductor's inductance and antenna patterns. However, it does not support the simulation of static magnetic field's effect on the inductor. The enhancement of tilted beam angle through the use of a magnet will be validated through measurement in section III.

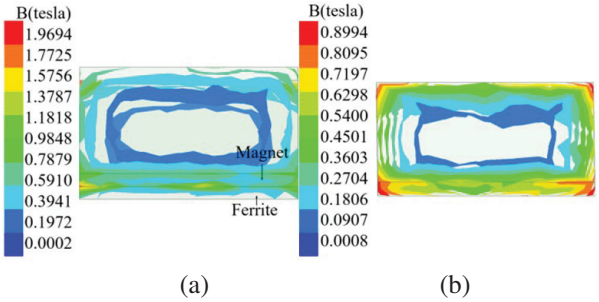


Fig. 8. Simulated B-field distribution of the magnet: (a) with ferrite and magnet and (b) with magnet.

III. FABRICATION AND MEASUREMENT

The proposed antenna was fabricated and measured to validate the simulation results. Figure 9 shows photographs of the fabricated antenna. Figure 10 shows measured and simulated S_{11} plots. As shown, measurements agree well with the simulation. The overlapped measured bandwidth of the five modes covers 2.39-2.42 GHz.

Figure 11 shows the measured and simulated radiation patterns. As shown, measurements agree well with the simulation except for a slight reduction of measured gain which the mounting of the ferrite slab might introduce.

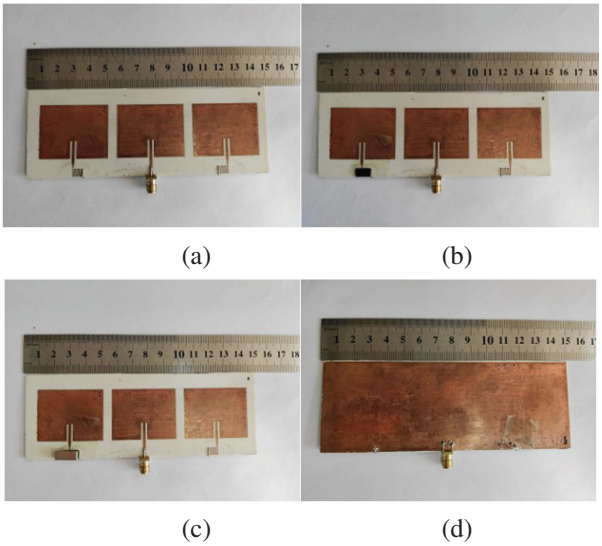


Fig. 9. Photographs of proposed antenna: (a) top layer of Mode 1, (b) top layer of Mode 3, (c) top layer of Mode 5, and (d) bottom layer.

The measured angles align closely with the simulated angles. Specifically, Mode 1 shows a steering angle of 0° with a measured gain of 6.82 dB, while Modes 2/3, with steering angles of 15°, both display nearly identical measured gains of 6.22 dB. In Modes 4 and 5, the deflection angles induced by the magnetic field attain 28°, with both configurations demonstrating nearly equivalent measured gains of 6.07 dB. There is little difference in the maximum gain of Modes 2 and 3 between measurement and simulation. The measured gain is a bit lower than the simulated one. This might be caused by the errors introduced in fabrication, such as dimension errors of ferrite slab and roughness at soldering points.

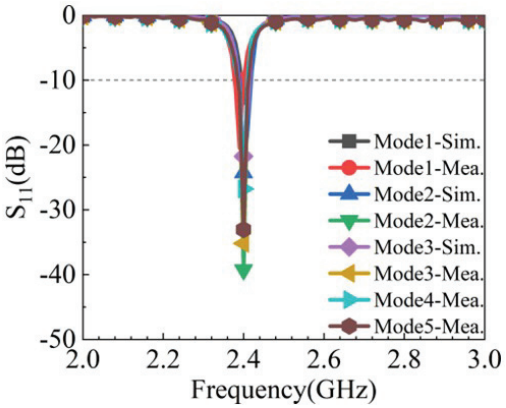


Fig. 10. Measured and simulated S_{11} of the proposed antenna.

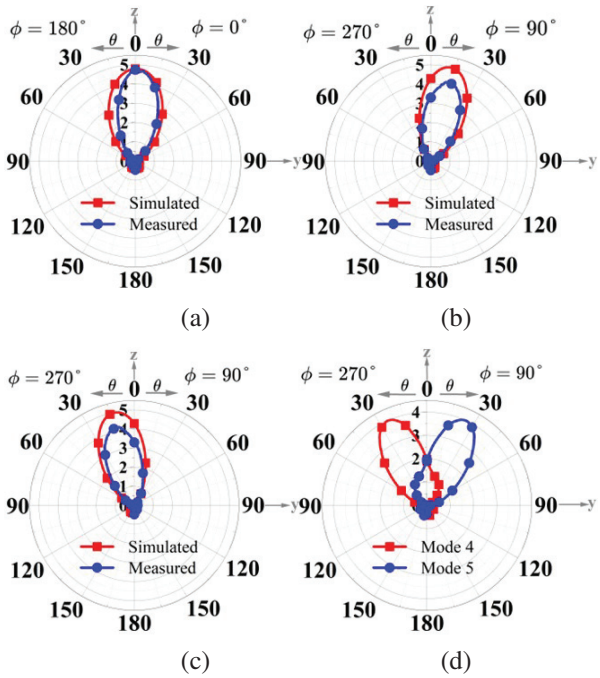


Fig. 11. Measured and simulated radiation patterns for different modes of yoz: (a) Mode 1, (b) Mode 2, (c) Mode 3, and (d) Modes 4 and 5.

IV. CONCLUSION

This paper presents a beam-reconfigurable antenna based on a patch antenna loaded with two distributed inductors with the same positive inductances. Adding a ferrite slab to one inductor changes its inductance to negative, resulting in a radiation beam tilted to the negative inductor's side. A larger tilted beam is obtained by applying static magnetic fields on the ferrite slab. The proposed antenna works at five beam states with the largest tilted angle of 28°. Compared with traditional beam reconfigurable antennas using switches or phase shifters, our design doesn't involve DC power or lumped components. It has the potential to find applications in several specialized fields, such as implantable devices with low/no DC power supply. It can also be used in a high-power system since the distributed inductors have higher power capacity than switches and the lumped components.

REFERENCES

[1] F. Ahmed, K. Singh, and K. P. Esselle, "State-of-the-art passive beam-steering antenna technologies: challenges and capabilities," *IEEE Access,* vol. 11, pp. 69101-69116, 2023.

[2] W. Li, Y. M. Wang, Y. Hei, B. Li, and X. Shi, "A compact low-profile reconfigurable metasurface antenna with polarization and pattern diversities," *IEEE Antennas Wirel. Propag. Lett*., vol. 20, no. 7, pp. 1170-1174, July 2021.

[3] S. Zhao, Z. Wang, and Y. Dong, "Pattern-reconfigurable antenna using low-profile electric and magnetic radiators," *IEEE Antennas Wirel. Propag. Lett*., vol. 22, no. 3, pp. 616-620, Mar. 2023.

[4] Y. Liu and Y. Wang, "Low-profile dual-polarized pattern-reconfigurable antenna with independent beam control capability in both polarizations for indoor application," *IEEE Antennas Wirel. Propag. Lett*., vol. 23, no. 4, pp. 1231-1235, Apr. 2024.

[5] Y. Mu, J. Han, D. Xia, H. Liu, and L. Li, "A self-decoupling wide-angle scanning pattern reconfigurable antenna array using parasitic comb structures," *IEEE Trans. Antennas Propag*., vol. 72, no. 2, pp. 1212-1223, Feb. 2024.

[6] K. Xue, H. Zhai, D. Yang, and Z. Wei, "A new pattern-reconfigurable antenna with the function of 360° wide-beam scanning and main beam direction finely-adjustable," *Microw. Opt. Technol. Lett*., vol. 60, pp. 2076-2081, 2018.

[7] Q. Liu, Z. Geng, R. Zhao, S. Li, Z. Yao, and W. H. Zong, "A wideband planar pattern reconfigurable

antenna for IEEE 802.11ac WLAN applications," *Int. J. RF Microw. Comput.-Aided Eng*., vol. 32, p. e23323, 2022.

[8] M. J. Kang, Y. S. Choi, and W. S. Lee, "Compact-size lightweight beam-reconfigurable ESPAR antenna with parasitic elements for UAV applications," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 39, no. 5, pp. 433-439, May 2024.

[9] G. R. Liu, C. Ju, Z. Li, Z. Yang, X. Qu, N. Liu, and W. Zong, "A low-profile wideband pattern reconfigurable antenna with metasurface," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 39, no. 10, pp. 868-875, Oct. 2024.

[10] Y. Yusuf and X. Gong, "A low-cost patch antenna phased array with analog beam steering using mutual coupling and reactive loading," *IEEE Antennas Wirel. Propag. Lett*., vol. 7, pp. 81-84, 2008.

[11] Q. H. Dang, S. J. Chen, N. Nguyen-Trong, and C. Fumeaux, "Multifunctional reconfigurable wearable textile antennas using coplanar reconfiguration modules," *IEEE Trans. Antennas Propag*., vol. 71, no. 5, pp. 3806-3815, May 2023.

[12] K. Y. Kapusuz, Y. Şen, M. Bulut, İ. Karadede, and U. Oğuz, "Low-profile scalable phased array antenna at Ku-band for mobile satellite communications," in *2016 IEEE International Symposium on PAST*, Waltham, MA, pp. 1-4, 2016.

[13] J. Hu and Z. C. Hao, "A compact polarization-reconfigurable and 2-D beam-switchable antenna using the spatial phase shift technique," *IEEE Trans. Antennas Propag*., vol. 66, no. 10, pp. 4986-4995, Oct. 2018.

[14] B. Y. Liu, H. W. Zhang, X. L. Tang, and H. Su, "RF integrated planar inductor fabricated on ferrite substrates," in *2006 International Conference on Communications, Circuits and Systems*, Guilin, China, 2006.

[15] C. Pardue, M. F. Bellaredj, A. K. Davis, and M. Swaminathan, "Miniaturization of planar packaged inductor using NiZn and low-cost screen-printing technique," in *2017 IEEE 67th ECTC,* Orlando, FL, 2017.

[16] T. Zheng, G. Xu, L. Luo, and T. Zheng, "A novel wafer level high Q planar inductor using Ni-Zn ferrite/BCB composite thick film," in *2016 17th ICEPT,* Wuhan, China, 2016.

[17] C. Sun, D. Qiu, P. Li, Z. Yao, P. Wang, and S. Li, "The enhancement of inductance and Q-factor in micro-magnetic inductors by self-biased FeCoB thin films," in *2023 IEEE 6th ICEICT*, Qingdao, China, pp. 1067-1069, 2023.

[18] D. S. Gardner, G. Schrom, F. Paillet, B. Jamieson, T. Karnik, and S. Borkar, "Review of on-chip inductor structures with magnetic films," *IEEE Trans. Magn*., vol. 45, no. 10, pp. 4760-4766, Oct. 2009.

[19] J. J. Lee, Y. K. Hong, S. Bae, J. Jalli, G. S. Abo, R. Syslo, B. C. Choi, and G. W. Donohoe, "High-quality factor Ni-Zn ferrite planar inductor," *IEEE Trans. Magn*., vol. 46, no. 6, pp. 2417-2420, June 2010.

[20] A. A. Adly, "Controlling linearity and permeability of iron core inductors using field orientation techniques," *IEEE Trans. Magn*., vol. 37, no. 4, pp. 2891-2893, July 2001.

[21] N. Ning, X. P. Li, J. Fan, W. C. Ng, Y. P. Xu, X. Qian, and H. L. Seet, "A tunable magnetic inductor," *IEEE Trans. Magn*., vol. 42, no. 5, pp. 1585-1590, May 2006.

[22] B. Assadsangabi, M. S. Mohamed Ali, and K. Takahata, "Planar variable inductor controlled by ferrofluid actuation," *IEEE Trans. Magn*., vol. 49, no. 4, pp. 1402-1406, Apr. 2013.

[23] J. Cui, H. Wang, L. Qu, and W. Qiao, "A tunable inductor based on a magnetic flux valve," in *2017 IEEE Energy Conversion Congress and Exposition (ECCE)*, Cincinnati, OH, pp. 3255-3259, 2017.

[24] B. M. F. Rahman, R. Divan, D. Rosenmann, T. Wang, Y. Peng, and G. Wang, "Application of sub-micrometer patterned permalloy thin film in tunable radio frequency inductors," *J. Appl. Phys*., vol. 117, p. 17C121, May 2015.

Meng Yin was born in Hebei, China. She is currently pursuing a master's degree at the School of Electronic Information, Qingdao University, China. Her primary research focus is on beam-reconfigurable antennas.

Cheng Ju received the doctor's degree from Beijing University of Posts and Telecommunications, China, in 2015. He is currently a master's supervisor at the School of Electronic Information, Qingdao University, China. His research interests include transceiver imbalance calibration and digital parallel processing implemented on FPGA.

Wei-Hua Zong received the doctor's degree from Xidian University, China, in 2004. She is currently an associate professor at the School of Electronic Information, Qingdao University, China. Her research interests include the design of implanted and wearable antennas, mobile phone antennas, base station antennas, and beam-scanning antennas.

Shandong Li received the doctor's degree from Nanjing University in 2003. He is currently a distinguished professor and doctoral supervisor at the School of Electronic Information, Qingdao University, China. His research interests include superconducting quantum interference devices and their applications in medicine, geophysics, and nondestructive testing; high-frequency soft magnetic materials and their integrated microwave devices; and weak magnetic field sensors and measurement systems based on magnetoelectric coupling effects.

Extended Network Parameters for Transmission Line Networks Subject to External Field Illumination

Mingwen Zhang[1], Chunguang Ma[1], Bicheng Zhang[1], and Yong Luo[1,2]

[1]School of Electronic Science and Engineering
University of Electronic Science and Technology of China, Chengdu, China
zmw95@foxmail.com, macg@uestc.edu.cn, 2020190503024@std.uestc.edu.cn, yluo@uestc.edu.cn

[2]State Key Laboratory of Electromagnetic Space Cognition and Intelligent Control Technology
Beijing, China

***Abstract* –** Microwave network parameters are prevalently used in the modeling of transmission lines. It can characterize the interconnection of ports and is supported by most SPICE software. However, traditional microwave network parameters cannot characterize the role of external fields; this is a common concern in electromagnetic compatibility or electromagnetic interference analysis. In this paper, external excitation is treated as an additional port in the circuit. An extended network parameter is proposed to model the transmission lines excited by the external field. The extended network parameters can be used easily in the SPICE solver to analyze responses on linear or nonlinear loads. The proposed method is suitable for evaluating the electromagnetic interference of complex transmission line networks or PCBs, and it can use the advantages of circuit solvers for tuning or optimization with less computational burden.

***Index Terms* –** Black-box modeling, electromagnetic coupling, network parameters, transmission line.

I. INTRODUCTION

With the rapidly growing complexity of the electromagnetic environment, the survivability of electronic devices or modules exposed to it should be valued. In a complex electromagnetic environment, those man-made high-power electromagnetic (HPEM) interference sources are undoubtedly the most threatening, such as high-power microwaves (HPM), which are defined as microwave sources that can be operated in 300 MHz to 300 GHz, and their power is more than megawatts [1, 2]. Transmission lines (TLs), such as micro-strip lines (MSLs), are widely employed in high-frequency circuits and systems. When exposed to electromagnetic radiation, an induced voltage or current is generated on the TL and transmitted to the load. Due to the widespread presence of nonlinear loads in susceptible systems, high-power signals may alter the operating state of these devices. The resulting effects include device-level rectification [3, 4] and system-level interference [5, 6], which can lead to improper functioning of the system. More seriously, when the induced voltage exceeds the capacity of the device, it may burn out [7–9]. Therefore, accurately and effectively predicting the induced voltage is of great significance for electromagnetic interference (EMI) evaluation, as it involves solving the field-line coupling problem.

Generally, field-to-line problems can be solved using multi-conductor transmission line theory (MTLT). However, it will not perform well at high frequencies because it is based on the quasi-TEM assumption [10, 11]. Full-wave and hybrid methods initially adopt techniques such as the finite-difference time-domain method [12], finite-element method [13], partial element equivalent circuit [14], and time-domain integral equation [15] to solve electromagnetic coupling issues. The response of terminations is then determined using corresponding circuit analysis methods, including transient convolution-based techniques, harmonic balance, and the envelope tracking method [16]. Although these methods can produce accurate results, they require significant resource consumption due to the need for discretizing both time and space. These approaches necessitate self-consistent solving of electromagnetic and electrical problems. Any changes in load or excitation waveform require a re-performance of the full-wave analysis. Consequently, if parameter tuning or statistical analysis is required, repetitive full-wave calculations can be laborious [17, 18]. Another approach involves decomposing the field-line coupling problem into a linear electromagnetic coupling problem and a circuit-solving problem through the Thevenin's equivalent network approach [19]. In [20], the response of linear and nonlinear loads inside a perforated metallic

Submitted On: January 14, 2025
Accepted On: April 2, 2025

https://doi.org/10.13052/2025.ACES.J.400408

enclosure to HPEM irradiation was investigated, and vector fitting (VF) is used to fit the Thevenin's equivalent impedance obtained from finite difference delay modeling.

Microwave port analysis is a fundamental technique in high-frequency circuit design and microwave engineering. It involves defining input and output ports to measure and analyze parameters such as scattering parameters (*S*-parameters), impedance parameters (*Y*-parameters), and admittance parameters (*Z*-parameters). These parameters are essential for modeling TLs such as cable bundles and PCB interconnects. Various techniques can integrate these network parameters into circuit solvers, whether in the time domain [21, 22] or frequency domain [23], enabling engineers to perform mixed-signal circuit analysis in complex systems such as high-density interconnect PCBs. However, traditional microwave network parameters do not account for external field effects, critical in EMI analysis. In [24], a hybrid *S*-parameter approach is proposed to address the field-to-line problem by decoupling the excitation signal into forced and modal waves using the generalized pencil of functions. In this paper, we introduce a similar but more general approach. Field-line coupling issues are decomposed into a linear electromagnetic problem and a circuit-solving problem through Thevenin's theorem. Electromagnetic results can be derived from measurement or full-wave simulation data, in contrast to [24], which does not require special decomposition of the external field. A black-box model involving $(N+1)$-port *Z*-parameters is directly derived, and the plane wave is introduced to the network as an additional port. The rest of this paper is organized as follows. In section II, the theory is developed by analyzing an *N*-port TL network illuminated by a plane wave. The $(N+1)$-port network parameters are derived, and its equivalent circuit is implemented using VF. In section III, we validate the concept with three examples: nonlinear loads, TL networks, and TL in an enclosure. Further details are discussed in section IV, and the conclusions are presented in section V.

II. THEORY ANALYSIS

Figure 1 (a) illustrates a typical scenario of an *N*-port TL network illuminated by a plane wave. Using Thevenin's theorem, this can be translated into the problem shown in Fig. 1 (b).

In Fig. 1 (b), ${V_m}^{OC}$ represents the open-circuit voltage at port m due to plane-wave illumination. Since coupling is a linear process, we can use the circuit shown in Fig. 2 to establish a connection between the external field and the open-circuit voltage. This circuit consists of an input impedance and a current-controlled voltage source. The voltage source V_{EMI} has the same waveform as the external excitation field. The input impedance at port m can be expressed as:

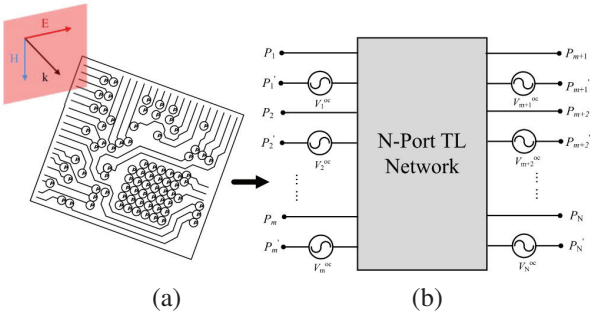


Fig. 1. TL networks subject to (a) plane wave illumination and (b) its equivalent representation.

$$Z\mathrm{in}_{\mathrm{m}}(s) = r \cdot \frac{V_{\mathrm{m}}^{\mathrm{oc}}(s)}{V_{\mathrm{EMI}}(s)}. \tag{1}$$

In (1), *r* represents the transresistance of the current-controlled voltage source, which is typically set to 1. However, to prevent Z_{in} from becoming numerically too small, the transresistance may be scaled to minimize numerical errors.

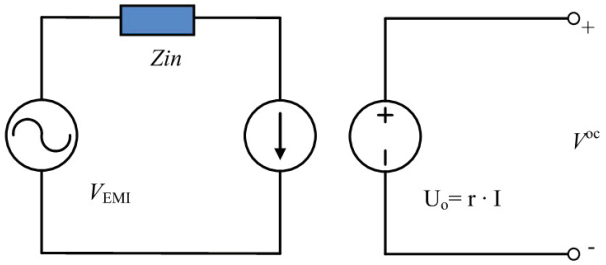


Fig. 2. Circuit connected V^{OC} and V_{EMI}.

Once we have established the connection between the V^{OC} in Fig. 1 (b) and the external field with the circuit shown in Fig. 2, we will then get the circuit as shown in Fig. 3.

Network parameter analysis is widely used in circuit analysis. In this paper, we will treat the external field as an additional port. The open-circuit impedance representation of Fig. 3 is expressed by:

$$V(s) = Z(s)I(s). \tag{2}$$

Then, the elements of the impedance matrix can be obtained with all the other ports left open-circuit:

$$Z_{\mathrm{ij}}(s) = \frac{V_{\mathrm{i}}(s)}{I_{\mathrm{j}}(s)}, I_{\mathrm{k}}(s) = 0, \forall k \neq j. \tag{3}$$

Suppose the intrinsic *Z*-parameter of the TL network under consideration is:

$$Z^{\mathrm{Intr.}} = \begin{bmatrix} Z_{1,1}^{\mathrm{Intr.}} & \cdots & Z_{1,\mathrm{N}}^{\mathrm{Intr.}} \\ \vdots & \ddots & \vdots \\ Z_{\mathrm{N},1}^{\mathrm{Intr.}} & \cdots & Z_{\mathrm{N,N}}^{\mathrm{Intr.}} \end{bmatrix}. \tag{4}$$

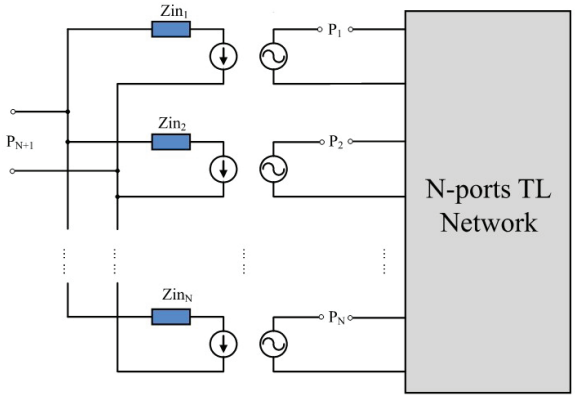


Fig. 3. Extended (N+1)-port network, and external fields are fed from the (N+1)-port.

(1) When $i \leq N$ and $j \leq N$, since the $(N+1)^{th}$ port is open circuit, the network parameters at these locations remain consistent with the intrinsic part.

(2) When $i=N+1$ and $j \leq N$, the external fields are not affected by the parameters of the TL network, so the parameters for these positions are set to 0:

$$Z_{\mathrm{i,j}}^{\mathrm{Extd.}} = 0. \tag{5}$$

(3) When $i \leq N$ and $j=N+1$, which means all the ports are open-circuited except for the external one, then $Z_{i,j}^{Extd}$ can be obtained through a simple circuit analysis:

$$Z_{\mathrm{i,N+1}}^{\mathrm{Extd.}} = \frac{V_{\mathrm{i}}}{I_{\mathrm{N+1}}} = -r_{\mathrm{i}} \frac{\mathrm{Zin_1 Zin_2 \cdots Zin_N}}{\mathrm{Zin_i}}$$
$$r_{\mathrm{i}} \mathrm{Zin_i} \sum_{m=1}^{N} \mathrm{Zin_m^{-1}}. \tag{6}$$

(4) When $i=N+1$ and $j=N+1$, then $Z_{N+1,N+1}^{\mathrm{Extd}}$ can be determined through a straightforward circuit analysis:

$$Z_{\mathrm{N+1,N+1}}^{\mathrm{Extd.}} = \frac{V_{\mathrm{N+1}}}{I_{\mathrm{N+1}}} = \mathrm{Zin_1 Zin_2 \cdots Zin_N}. \tag{7}$$

For N-port network illuminated by plane-wave, the extended Z-parameters can be expressed as:

$$Z^{\mathrm{Extd.}} = \begin{bmatrix} & & -\frac{r_1}{\mathrm{Zin_1} \sum_{m=1}^{N} \mathrm{Zin_m^{-1}}} \\ & Z^{\mathrm{Intr.}} & \vdots \\ & & -\frac{r_{\mathrm{N}}}{\mathrm{Zin_N} \sum_{m=1}^{N} \mathrm{Zin_m^{-1}}} \\ 0 \quad 0 \quad 0 & & \frac{1}{\sum_{m=1}^{N} \mathrm{Zin_m^{-1}}} \end{bmatrix}$$
$$= \begin{bmatrix} Z_{1,1}^{\mathrm{Intr.}} & \cdots & Z_{\mathrm{N,1}}^{\mathrm{Intr.}} & -\frac{r_1}{\mathrm{Zin_1} \sum_{m=1}^{N} \mathrm{Zin_m^{-1}}} \\ \vdots & \ddots & \vdots & \vdots \\ Z_{1,\mathrm{N}}^{\mathrm{Intr.}} & \cdots & Z_{\mathrm{N,N}}^{\mathrm{Intr.}} & -\frac{r_{\mathrm{N}}}{\mathrm{Zin_N} \sum_{m=1}^{N} \mathrm{Zin_m^{-1}}} \\ 0 & \cdots & 0 & \frac{1}{\sum_{m=1}^{N} \mathrm{Zin_m^{-1}}} \end{bmatrix}. \tag{8}$$

When N=2, the network extends to a three-port network, and the Z-parameter is written as:

$$Z^{\mathrm{Extd.}} = \begin{bmatrix} Z_{1,1}^{\mathrm{Intr.}} & Z_{1,2}^{\mathrm{Intr.}} & -r_1 \frac{\mathrm{Zin_2}}{\mathrm{Zin_1}+\mathrm{Zin_2}} \\ Z_{2,1}^{\mathrm{Intr.}} & Z_{2,2}^{\mathrm{Intr.}} & -r_2 \frac{\mathrm{Zin_1}}{\mathrm{Zin_1}+\mathrm{Zin_2}} \\ 0 & 0 & \frac{\mathrm{Zin_1 Zin_2}}{\mathrm{Zin_1}+\mathrm{Zin_2}} \end{bmatrix}. \tag{9}$$

So far, we have derived the extended Z-parameter. This can be translated into other network parameters, such as the Y-parameter or the S-parameter, which are typically presented in textual form. If the electromagnetic designer only needs to provide the results of the field analysis to the circuit designer, their task is considered complete at that point. As a black-box modeling method, most circuit solvers can identify network parameters, but they do not necessarily support the use of network parameters for circuit analysis directly, so it is necessary to introduce its circuit implementation.

As a robust numerical fitting method, VF employs N-order rational functions to fit transfer functions in the frequency domain [25–27], which denotes:

$$F(s) = \sum_{k=1}^{N} \frac{r_{\mathrm{k}}}{s - p_{\mathrm{k}}} + se + d, \tag{10}$$

where $s=j\omega$ signifies the complex frequency; r_k and p_k denotes the k-th residue and pole, respectively, with both being real or conjugate complex numbers; e and d represent linear and constant coefficients, respectively.

When residues and poles appear as conjugate complex numbers, the transfer function can be expressed as:

$$f(s) = \frac{r}{s-p} + \frac{r}{s-p^*}$$
$$= \frac{2\Re(r)s - 2(\Re(p)\Re(r) + \Im(p)\Im(r))}{s^2 - 2\Re(p)s + (\Re(p)^2 - \Im(p)^2)} \tag{11}$$

where $\Re$ and $\Im$ denote the real and imaginary parts of a complex number, respectively. Figure 4 illustrates a circuit implementation of the transfer function expressed in the impedance domain.

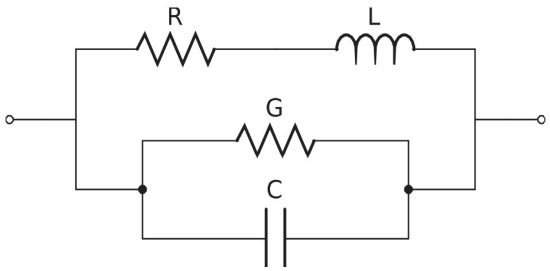


Fig. 4. Equivalent circuit for conjugated residue and pole.

The impedance in Fig. 4 is given by:

$$Z(s) = \frac{1}{\frac{1}{R+sL} + sC + G} = \frac{\frac{1}{C}s + \frac{R}{LC}}{s^2 + \left(\frac{R}{L} + \frac{G}{C}\right)s + \frac{1+GR}{LC}}. \tag{12}$$

By comparing (11) and (12), the value of the lumped elements in Fig. 4 can be determined:

$$\begin{aligned} C &= \frac{1}{2\Re e(r)} \\ G &= -\frac{\Im(p)\Im(r)+\Re(p)\Re(r)}{2\Re(r)^2} \\ L &= \frac{2\Re(r)^3}{\Im(p)^2(\Im(r)^2-\Re(r)^2)} \\ R &= \frac{2\Re(r)^2(\Im(p)\Im(r)-\Re(p)\Re(r))}{\Im(p)^2(\Im(r)^2-\Re(r)^2)}. \end{aligned} \tag{13}$$

Although VF can be used to model multi-port networks, we recommend fitting the TL network and Z_{in} in Fig. 3 separately. This approach is more straightforward and allows for equivalent circuit modeling of TLs through alternative methods. For a more comprehensive understanding of TL network modeling methods, readers can refer to [28, 29].

A brief description of the $(N+1)$-port black-box modeling is:

(1) Obtain the transfer function from full-wave simulation or measurement.

(2) Building an extended network incorporates external excitation interaction with the process in (2) to (10).

(3) Fitting the network parameters uses VF, while implementing the circuit employs lumped components.

(4) Performing the field-line coupling analysis in circuit solver.

III. VALIDATION STUDIES

A. Microstrip line terminal with nonlinear load

The first example involves a two-port network, with the detailed geometry shown in Fig. 5. It consists of a single MSL with a width of 1.5 mm, situated on a substrate with a dielectric constant of 2.2 and a thickness of 0.8 mm. The overall dimensions of the board are 80×60 mm, and the thickness of the conductor layer is 35 μm. A plane wave propagates along the $-z$-axis, and the electric field is oriented along the x-axis.

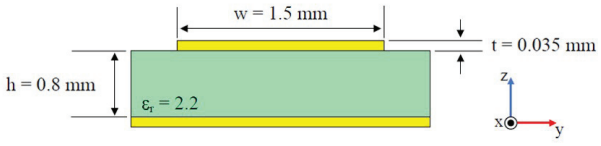


Fig. 5. MSL section geometry diagram.

The excitation waveform is double exponential pulses, which is defined as the early time high-altitude electromagnetic pulse waveform in IEC 61000-2-9 [30], which is denoted as:

$$s(t) = E_0 k \left(e^{-\alpha t} - e^{-\beta t}\right). \tag{14}$$

In this study, the parameters are set as E_0 =50 kV/m, k=1.3, α =4×10^7 s^{-1}, and β=6×10^8 s^{-1}, with the waveform shown in Fig. 6.

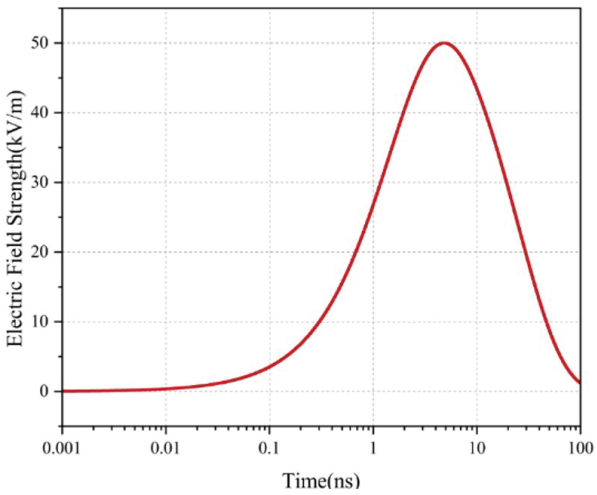


Fig. 6. High-altitude electromagnetic pulse waveform in time domain.

To obtain the extended network parameters for this study, full-wave simulations were conducted to determine each element in the matrix. Subsequently, the input impedance Z_{in} and TL networks were fitted from 10 MHz to 10 GHz using VF. Figure 7 illustrates the fitting results of Z_{in}, while Fig. 8 presents the full-wave calculations and the fitting results of the S-parameters of the TL, demonstrating good agreement between them.

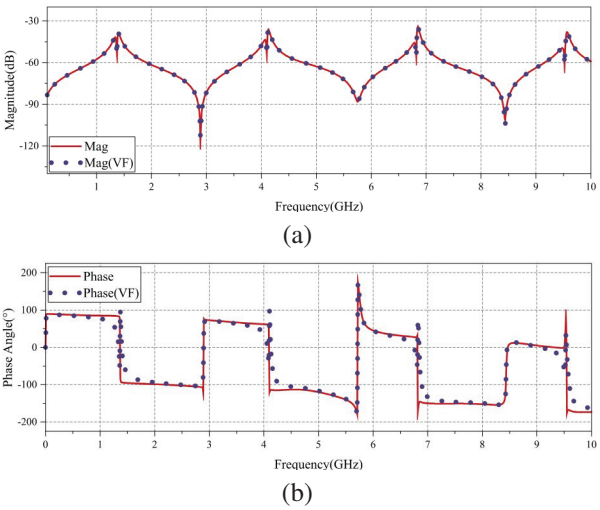
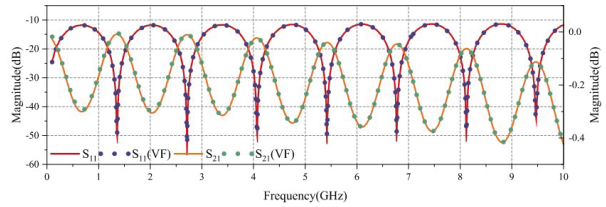
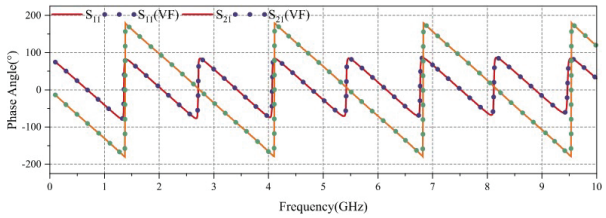


Fig. 7. (a) Amplitude and (b) phase of the input impedance (Z_{in}) of port 1, with the dashed line representing the VF results.

After the network parameters are determined and the circuit is implemented, their effectiveness is verified. In the validation example, port 1 is connected to a 50Ω resistor, while port 2 is connected to a nonlinear diode,



(a)



(b)

Fig. 8. (a) Amplitude and (b) phase response of the MSL corresponding to the geometry shown in Fig. 5. The dashed lines represent the vector fitting (VF) results.

with its equivalent circuit depicted in Fig. 9, and the nonlinear $I-V$ equation of the diode is:

$$I = \begin{cases} 0, & U < 0 \\ I_0(\exp^{\frac{e}{kT}U} - 1), & U \geq 0 \end{cases}, \tag{15}$$

where $I_0 = 5\times10^{-5}$A signifies the reverse saturation current, T=300 K, e denotes the elementary charge, and k represents the Boltzmann constant.

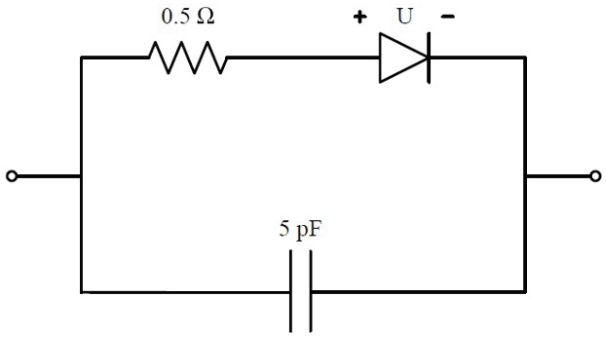
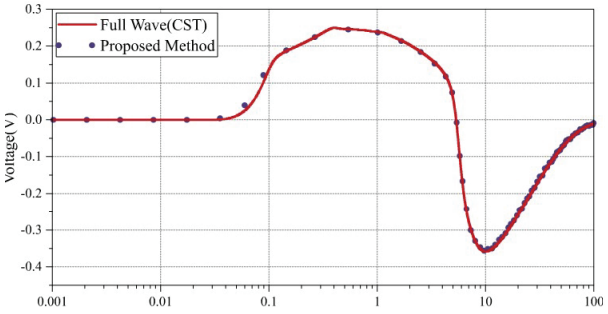


Fig. 9. Equivalent circuit of the diode.

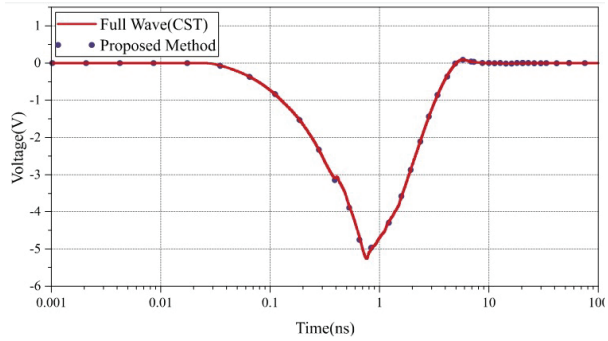
Figure 10 shows the calculation results of the full-wave and extended network parameters, and the two agree well. Note that the peak occurrence time of the induced voltage does not coincide with that of the excitation voltage.

B. MSLs with multi-ports

In general, TL networks are more of a concern in the EMC community or industry. In this example, a network with five-port is studied. Figure 11 shows the detailed geometry, four-port are connected between the MSL termination and the reference ground plane, with one-port



(a)



(b)

Fig. 10. Induced voltage on (a) resistor and (b) diode.

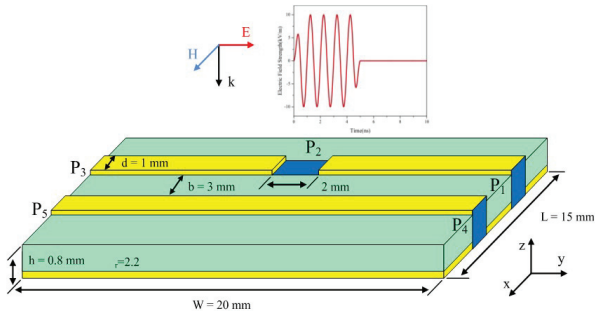


Fig. 11. Pair of microstrip lines exposed to plane-wave illumination.

spanning MSLs. This configuration is more realistic and closely resembles a typical PCB layout.

The impressed electric field, as shown in Fig. 11, consists of a plane wave incident normal to the ground plane, with the electric field aligned along the TL axis. The incident waveform is considered as a narrowband HPM [31], illustrated in (16):

$$E_{\text{inc}}(t) = \begin{cases} E_0 \frac{t}{t_{\text{r}}} \sin(2\pi f_0 t) & 0 < t < t_{\text{r}} \\ E_0 \sin(2\pi f_0 t) & t_r < t < t_r + \tau \\ E_0 \left(\frac{\tau + t_{\text{r}} + t_{\text{f}}}{t_{\text{f}}} - \frac{t}{t_{\text{r}}} \right) & t_{\text{r}} + \tau < t < t_{\text{r}} + t_{\text{f}} + \tau \end{cases}, \tag{16}$$

where E_0 denotes the peak electric field strength; τ signifies the pulse width; t_r and t_f represent the rise time

and fall time, respectively; f_0 is the carrier frequency of the HPM pulse. In the following example, E_0 =10 kV/m, τ =4 ns, t_r =t_f =0.5 ns.

Once again, we need to extract the extended network parameters and implement the circuit. For the sake of brevity, we will not present this section in detail. Instead, we will show the results obtained from both full-wave simulations and this method when the port is terminated with various loads. Figure 13 shows the induced voltage for each load in the circuit configuration shown in Fig. 12. It can be observed that the full-wave calculations are consistent with the method presented in this paper. The value of the induced voltage depends on the type and value of the load. Due to the presence of the energy storage element, the induced voltage does not immediately return to zero after the pulse ends, as discussed in detail in [3].

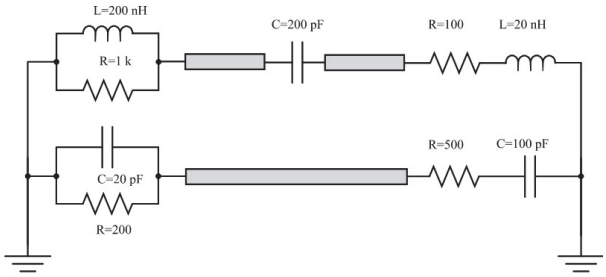
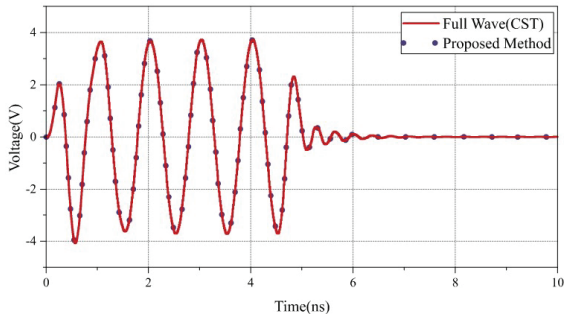
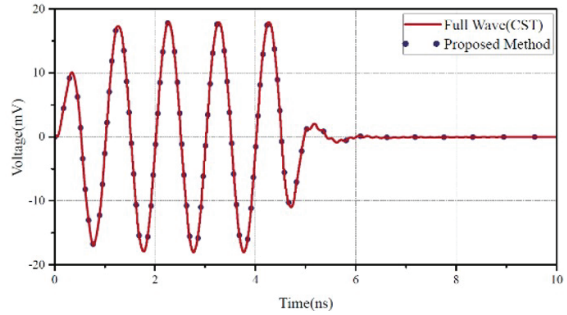


Fig. 12. Schematic of port configuration.

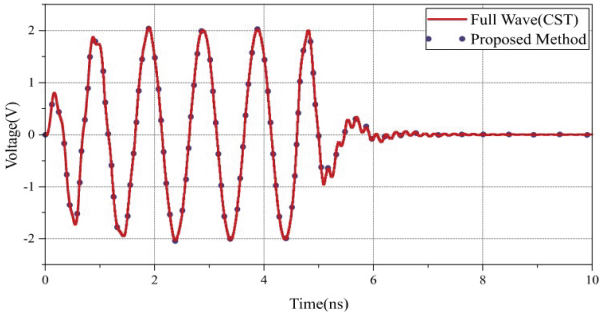


(a)

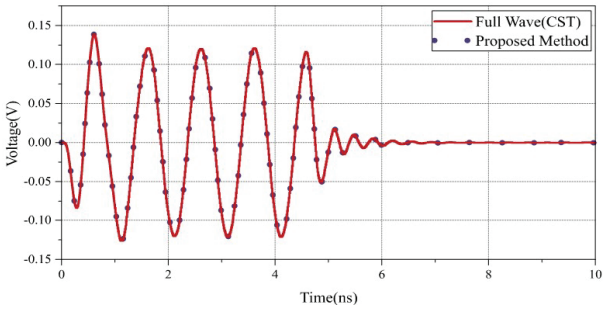


(b)

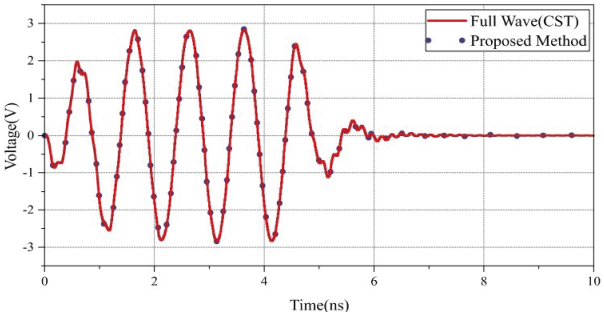
Fig. 13. *Continued.*



(c)



(d)



(e)

Fig. 13. Induced voltage at different loads: (a-e) represent the induced voltages on ports 1-5, respectively.

C. MSL in perforated metallic enclosure

To resist the threat of HPEM, designers often use metal enclosures to protect the internal circuitry, however, due to the need for ventilation or wiring, there will be some slots or apertures in the cavity, which degrades the shielding ability of the cavity. Therefore, we need to evaluate the immunity of circuits in a perforated metallic enclosure. In this example, a two-port TL located in a perforated enclosure as shown in Fig. 14 is considered.

In this example, the enclosure dimensions $L\times W\times H$ are 22×17×5 mm, and the dimensions of the perforated rectangular aperture are $cx\times cy$=5×1 mm. The substrate measures $a\times b\times c$=20×15×0.8 mm with a dielectric constant of 2.2. The length and width of the MSL are 20 mm and 1 mm, respectively. The TL is then illuminated by a plane wave with the propagation direction

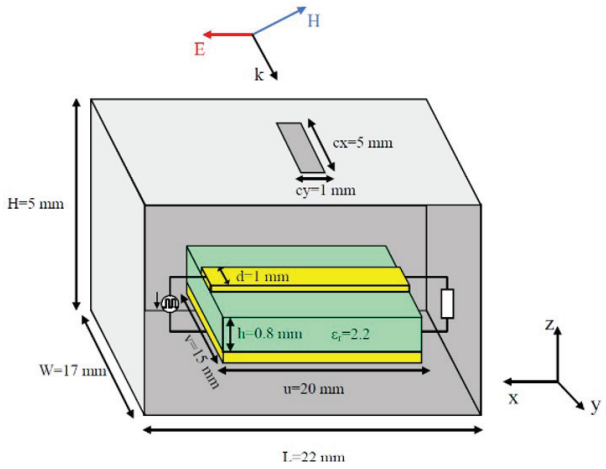


Fig. 14. Schematic of the two-port circuit illuminated by IEMI source.

k=(0, 0.154, -0.988), which deviates slightly from the $-z$-axis. The waveform is a Gaussian waveform, as shown in Fig. 15.

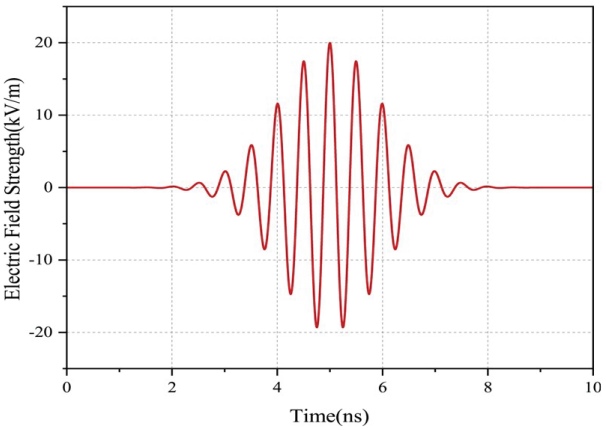


Fig. 15. Time domain representation of incident Gaussian waveform.

The expression for the waveform is given as:

$$E_{\text{inc}}(t) = E_0 \cos(2\pi f_0 t) e^{-\left(\frac{1}{2\sigma^2}(t-t_{\text{d}})\right)^2}, \qquad (17)$$

where E_0 =20 kV/m, σ=1 ns, f_0 =2 GHz and t_d =5 ns. Port 1 is connected to a 2 GHz rectangular pulse current source, and port 2 is connected to a 50 Ω resistor. It can be regarded as a clock signal in a digital system. Figure 16 shows the signal on the resistor. The method described in this paper is consistent with the full-wave results. Additionally, we observe that the signal undergoes significant distortion under the illumination of an external field.

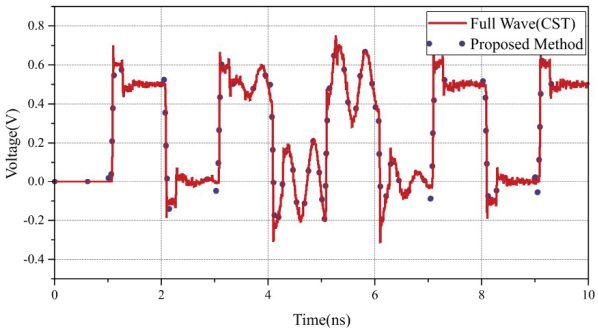


Fig. 16. Induced voltage on the resistor.

IV. DISCUSSION

In this work, we have developed an extended network parameter to model TL under external field illumination. The following discussion highlights several key issues that need to be addressed.

(1) In general, the coverage of network parameters should cover the frequency band we are interested in. In the case of the presence of nonlinear loads, the influence of harmonics should also be considered. The extraction range of network parameters should cover as many harmonics as possible.

(2) The simulation may encounter convergence issues due to the model being established within a finite bandwidth, which can result in non-causal responses. Relevant solutions can be found in [19]. Additionally, circuit convergence can be affected by the violation of passivity in the macromodel. Thus, during vector fitting, the passivity of the model should be ensured.

(3) In this paper, we have introduced the Z-parameters, through which we observe the clear influence of external fields. However, S-parameters are supported by a broader range of circuit solvers. Certain simulation software, like ADS, facilitates direct frequency domain analysis using S-parameters. In some instances, using a circuit file, such as a SPICE netlist, may be more suitable for compatibility reasons, as it is widely supported.

(4) The last issue is its applicability. Although this paper mainly focuses on MSLs, the theory can also be generalized to other types of TLs, such as cables. In addition, in the verification section of the paper, electromagnetic data primarily originates from full-wave simulations. As mentioned earlier, electromagnetic data can also be obtained through measurements using a network analyzer test in an electromagnetic anechoic chamber or a transverse electromagnetic cell.

V. CONCLUSION

In this study, extended network parameters were proposed and applied for the analysis of TLs under

external field irradiation. We demonstrated their effectiveness through three examples, involving both shielded and unshielded scenarios, linear and nonlinear loads, as well as single TL and TL networks. The simulation results from the circuit were consistent with the full-wave simulation results but required less computation time. The extended network parameters incorporate the influence of external fields, enabling the analysis of field-line coupling problems within the circuit solver without the need for repeated field calculations. Furthermore, this method facilitates a direct connection between the electromagnetic solver and the circuit solver, allowing the export of electromagnetic calculation results to the circuit solver. This approach is particularly suitable for large-scale projects that necessitate collaborative design efforts among multiple designers in the realm of electromagnetic and circuits.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 61921002.

REFERENCES

[1] D. V. Giri, R. Hoad, and F. Sabath, "Implications of high-power electromagnetic (HPEM) environments on electronics," *IEEE Electromagnetic Compatibility Magazine*, vol. 9, no. 2, pp. 37-44, 2020.

[2] W. A. Radasky and R. Hoad, "Recent developments in high power EM(HPEM) standards with emphasis on high altitude electromagnetic pulse (HEMP) and intentional electromagnetic interference (IEMI)," *IEEE Letters on Electromagnetic Compatibility Practice and Applications*, vol. 2, no. 3, pp. 62-66, 2020.

[3] R. Michels, M. Kreitlow, A. Bausen, C. Dietrich, and F. Gronwald, "Modeling and verification of a parasitic nonlinear energy storage effect due to high-power electromagnetic excitation," *IEEE Transactions on Electromagnetic Compatibility*, vol. 62, no. 6, pp. 2468-2475, 2020.

[4] C. Pouant, J. Raoult, and P. Hoffmann, "Large domain validity of MOSFET microwave-rectification response," in *2015 10th International Workshop on the Electromagnetic Compatibility of Integrated Circuits (EMC Compo),* pp. 232-237, 10-13 Nov. 2015.

[5] T. Dubois, J. J. Laurin, J. Raoult, and S. Jarrix, "Effect of low and high-power continuous wave electromagnetic interference on a microwave oscillator system: From VCO to PLL to QPSK receiver," *IEEE Transactions on Electromagnetic Compatibility*, vol. 56, no. 2, pp. 286-293, 2014.

[6] Y. Bayram, J. L. Volakis, S. K. Myoung, S. J. Doo, and P. Roblin, "High power EMI on RF amplifier and digital modulation schemes," *IEEE Transactions on Electromagnetic Compatibility*, vol. 50, no. 4, pp. 849-860, 2008.

[7] J. E. Baek, Y. M. Cho, and K. C. Ko, "Damage modeling of a low-noise amplifier in an RF front-end induced by a high-power electromagnetic pulse," *IEEE Transactions on Plasma Science*, vol. 45, no. 5, pp. 798-804, 2017.

[8] Q. S. Liang, C. C. Chai, H. Wu, Y. Q. Liu, F. X. Li, and Y. T. Yang, "Mechanism analysis and thermal damage prediction of high-power microwave radiated CMOS circuits," *IEEE Transactions on Device and Materials Reliability*, vol. 21, no. 3, pp. 444-451, 2021.

[9] M. G. Backstrom and K. G. Lovstrand, "Susceptibility of electronic systems to high-power microwaves: Summary of test experience," *IEEE Transactions on Electromagnetic Compatibility*, vol. 46, no. 3, pp. 396-403, 2004.

[10] C. R. Paul, "Frequency response of multiconductor transmission lines illuminated by an electromagnetic field," *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-18, no. 4, pp. 183-190, 1976.

[11] F. Rachidi, "A review of field-to-transmission line coupling models with special emphasis to lightning-induced voltages on overhead lines," *IEEE Transactions on Electromagnetic Compatibility*, vol. 54, no. 4, pp. 898-911, 2012.

[12] Q. F. Liu, C. Ni, H. Q. Zhang, and W. Y. Yin, "Lumped-network FDTD method for simulating transient responses of RF amplifiers excited by intentional electromagnetic interference signals," *IEEE Transactions on Electromagnetic Compatibility*, vol. 63, no. 5, pp. 1512-1521, 2021.

[13] W. Rui and J. Jian Ming, "Incorporation of multiport lumped networks into the hybrid time domain finite-element analysis," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 8, pp. 2030-2037, 2009.

[14] X. Wang, A. Huang, W. Zhang, R. Yazdani, D.-H. Kim, and T. Enomoto, "Methodology for analyzing coupling mechanisms in RFI problems based on PEEC," *IEEE Transactions on Electromagnetic Compatibility*, vol. 65, no. 3, pp. 761-769, 2023.

[15] H. H. Zhang, L. J. Jiang, and H. M. Yao, "Embedding the behavior macromodel into TDIE for transient field-circuit simulations," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 7, pp. 3233-3238, 2016.

[16] T. Wendt, M. D. Stefano, C. Yang, S. Grivet-Talocia, and C. Schuster, "Iteration dependent waveform relaxation for hybrid field nonlinear circuit problems," *IEEE Transactions on*

Electromagnetic Compatibility, vol. 64, no. 4, pp. 1124-1139, 2022.

[17] T. Liang, G. Spadacini, F. Grassi, and S. A. Pignari, "Worst-case scenarios of radiated-susceptibility effects in a multiport system subject to intentional electromagnetic interference," *IEEE Access*, vol. 7, pp. 76500-76512, 2019.

[18] F. Vanhee, D. Pissoort, J. Catrysse, G. A. E. Vandenbosch, and G. G. E. Gielen, "Efficient reciprocity-based algorithm to predict worst case induced disturbances on multiconductor transmission lines due to incoming plane waves," *IEEE Transactions on Electromagnetic Compatibility*, vol. 55, no. 1, pp. 208-216, 2013.

[19] J. T. Williams, L. D. Bacon, M. J. Walker, and E. C. Zeek, "A robust approach for the analysis of EMI/EMC problems with nonlinear circuit loads," *IEEE Transactions on Electromagnetic Compatibility*, vol. 57, no. 4, pp. 680-687, 2015.

[20] A. Kalantarnia, A. Keshtkar, and A. Ghorbani, "Predicting the effects of HPEM radiation on a transmission line terminated with linear/nonlinear load in perforated metallic enclosure using FDDM/VF," *IEEE Transactions on Plasma Science*, vol. 48, no. 3, pp. 669-675, 2020.

[21] D. Winklestein, M. B. Steer, and R. Pomerleau, "Simulation of arbitrary transmission line networks with nonlinear terminations," *IEEE Transactions on Circuits and Systems*, vol. 38, no. 4, pp. 418-422, 1991.

[22] J. E. Schutt Aine and R. Mittra, "Scattering parameter transient analysis of transmission lines loaded with nonlinear terminations," *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, no. 3, pp. 529-536, 1988.

[23] K. S. Kundert and A. Sangiovanni Vincentelli, "Simulation of nonlinear circuits in the frequency domain," *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, vol. 5, no. 4, pp. 521-535, 1986.

[24] Y. Bayram and J. L. Volakis, "Hybrid S-parameters for transmission line networks with linear/nonlinear load terminations subject to arbitrary excitations," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 5, pp. 941-950, May 2007.

[25] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Transactions on Power Delivery*, vol. 14, no. 3, pp. 1052-1061, 1999.

[26] B. Gustavsen, "Improving the pole relocating properties of vector fitting," *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1587-1592, 2006.

[27] B. Salarieh and H. M. J. De Silva, "Review and comparison of frequency-domain curve-fitting techniques: Vector fitting, frequency-partitioning fitting, matrix pencil method and loewner matrix," *Electric Power Systems Research*, vol. 196, 2021.

[28] B. Gustavsen and A. Semlyen, "Admittance-based modeling of transmission lines by a folded line equivalent," *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 231-239, 2009.

[29] A. Dounavis, R. Achar, and M. Nakhla, "A general class of passive macromodels for lossy multiconductor transmission lines," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 10, pp. 1686-1696, 2001.

[30] IEC 61000-2-9, Electromagnetic Compatibility (EMC) – Part 2-9: Description of HEMP Environment — Radiated Disturbance, 1996.

[31] J. Benford, J. A. Swegle, and E. Schamiloglu, *High Power Microwaves*. Boca Raton, Florida: CRC Press, 2015.

Mingwen Zhang received the B.S. and M.S. degrees in electronic science and technology from University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2017 and 2020, respectively. He is pursuing the Ph.D. degree with UESTC, Chengdu, China. His research interests include electromagnetic interference and microelectronic reliability.

Chunguang Ma received the B.S. degree in physics from Anqing Normal University, Anqing, China, in 2008, and the Ph.D. degree in plasma physics from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2015. He was with the College of Engineering, University of Connecticut, Storrs, CT, USA, as a Visiting Researcher, from 2013 to 2014. He was with the School of Physical Electronics, UESTC, as a Lecturer, from 2015 to 2018. He is with the School of Electronic Science and Engineering, UESTC, as an Associate Researcher. His research interests include electromagnetic compatibility and interference, microwave and millimeter wave devices, circuits and systems, and ultra-wideband radar systems.

Bicheng Zhang received the B.S. degree in Electronic and Information Engineering from University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2024. He is pursuing the M.S. degree with UESTC. His research interests include electromagnetic interference and electromagnetic coupling.

Yong Luo received the Ph.D. degree in physical electronics from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2003. Since 1988, he has been with the Optoelectronics Center, UESTC, where he has been a Professor with the School of Physical Electronic, since 1999. He is also a researcher of with the Laboratory of Electromagnetic Space Cognition and Intelligent Control, Beijing, China. His research interests include gyrotron traveling-wave tube, high-power millimeter-wave technology and its application.

Cross-Coupled Wide Stopband SIW Bandpass Filter Loading Snake-shaped Slot

Yawen He[1], Zhonghua Ma[1*], Mengnan Wang[1], Jingyao Huang[1], and Yanfeng Jiang[2*]

[1]School of Marine Information Engineering
Jimei University, Xiamen 361021, China
yw_he@jmu.edu.cn, mzhxm@jmu.edu.cn, 202211810012@jmu.edu.cn, 202321303073@jmu.edu.cn

[2]College of IoT Engineering
Jiangnan University, Wuxi 214122, China
jiangyf@jiangnan.edu.cn
*Corresponding authors: Zhonghua Ma (mzhxm@jmu.edu.cn); Yanfeng Jiang (jiangyf@jiangnan.edu.cn)

Abstract – In this paper, a novel fourth-order bandpass filter is proposed and fulfilled, which is based on a cross-coupling structure achieved by etching snake-shaped slots between vias of adjacent SIW cavities. The gap of the vias' wall between adjacent SIW cavities is used to achieve coupling between the two cavities. The length of the coupling gap can be adjusted to change the resonant frequency of TE_{101} mode to TE_{201} mode. The offset distance between the ports and the centerline is disclosed as a key parameter to suppress undesired modes, achieving wide stopband characteristics. The snake-shaped slot is loaded on the vias' wall between SIW cavities for the enhancement of the electrical coupling, which produces a pair of transmission zeros (TZs) on either side of the passband, enhancing frequency selectivity. The filter design is validated through simulation, fabrication, and measurement of a fourth-order SIW filter with TZs. It is verified that the designed filter shows promising characteristics, with a center frequency of 24.1 GHz, an insertion loss of 0.52 dB, and a return loss exceeding 11.8 dB. It is observed that two TZs appear at 22.82 GHz and 25.15 GHz, respectively, each exhibiting a suppression level better than 20 dB. The designed filter shows potential applications in microwave and radio frequency circuits, as well as in wireless communication systems.

Index Terms – Bandpass filter, cross-coupling, transmission zeros, wide stopband.

I. INTRODUCTION

The rapid development of wireless communication systems accelerates the demand for compact, high-performance, multi-band filters. Modern communication systems, such as 5G, 6G, Internet of Things (IoTs), and satellite communication, are associated with increasingly high-performance requirements for the filter. In particular, for high-frequency signal transmission, filters with low insertion loss, high selectivity, and good out-of-band rejection play a key role in the system. Substrate Integrated Waveguide (SIW) has emerged as an ideal choice for designing high-end filters due to its advantages such as low cost, low loss, small size, high Q-factor, and seamless integration with other planar microwave circuits [1–3]. Recently, various SIW filters with different topologies and filtering characteristics have been proposed [4, 5]. To achieve sharp frequency selectivity, transmission zeros (TZs) must be inserted on both sides of the filter's passband [6–8]. Extensive research has been conducted in the community on the SIW filter, exploring different geometries and coupling mechanisms to realize TZs. However, limitations are still faced in balancing high selectivity and low insertion loss.

In current methods, additional structures are typically etched on the top layer of the SIW cavity to form mixed coupling of electric and magnetic fields, thereby introducing TZs on both sides of the passband in bandpass filters [9]. However, the adoption of these hybrid SIW structures may deteriorate the integrity and the shielding effectiveness of SIW cavities, potentially increasing radiation losses. Although electrical coupling is provided by the slot-line structure of the cascaded SIW cavities, and it can also be used as resonators, the circuit size of the filter is reduced simultaneously [10–16]. However, it should be noted that the design of these slot-line structures is very complex and could seriously interfere with the surface current distribution.

To overcome these challenges, attempts have been made to achieve multi-mode SIW filters by introducing metallized via perturbations [17–20]. However, parasitic capacitance and inductance are introduced inevitably by the vias. These parasitic effects would cause changes in the distribution of the electromagnetic fields, which

Submitted On: December 6, 2024
Accepted On: April 2, 2025

https://doi.org/10.13052/2025.ACES.J.400409

affect the frequency response of the filter. The resonance frequency would be shifted, and the bandwidth would be narrowed due to these parasitic effects. The influences would be more severe at microwave and radio frequency. In multi-layer SIW filters, the vertically stacked circuit structure shows more flexible options for implementing electrical coupling [21]. However, production costs are increased, and circuit layout flexibility is limited by the complexity of designing and manufacturing multilayer structures. Therefore, a solution that can provide effective out-of-band suppression and low insertion loss is urgently needed to keep the filter's shielding effectiveness and to simplify its multi-layer structure.

The design of SIW-based filters generally faces several challenges, including the balance between low insertion loss and high manufacturing complexity, as well as the impact of parasitic effects on performance. To address these issues, a novel filter design is proposed in this paper, which combines serpentine slots with optimized via configurations. The serpentine slot design is simple and maintains continuous working currents on the surface, effectively enhancing the filter's performance. By introducing the serpentine slot design based on a single-layer PCB structure, cross-coupling is successfully achieved, and TZs are introduced on both sides of the passband, effectively suppressing higher-order frequencies.

In [22], a design approach is reported where the coupling slot width of a SIW filter is adjusted to achieve flexible coupling. Through this method, a TZ is introduced in a half-mode SIW filter, and a bandpass filter is successfully constructed by integrating SIW and half-mode SIW (HMSIW) structures. However, while this design optimizes in-band filter performance, it does not effectively suppress higher-order modes, leading to spectral leakage and undesired resonances at higher frequencies. In modern millimeter-wave systems, unsuppressed higher-order modes cause severe interference in densely packed frequency bands, degrading signal integrity and system performance. This limitation underscores the need for novel filter designs that enhance both selectivity and out-of-band rejection.

To address this issue, this study proposes an innovative strategy that combines a serpentine slot design with an optimized via configuration. This approach successfully introduces multiple TZs, substantially improving out-of-band rejection on both sides of the passband. Additionally, by optimizing via spacing and serpentine slot geometry, the influence of higher-order modes is effectively mitigated. Unlike the approach in [22], which does not account for higher-order mode suppression, the proposed design exhibits significantly improved attenuation performance, particularly in the millimeter-wave range. This contribution not only addresses the shortcomings of prior SIW-based designs but also provides a robust and practical foundation for the development of next-generation high-performance filters.

In the proposed design, the manufacturing process is simplified with a single layer, with the merits of low cost and enhanced flexibility of circuit layout. A coupling window is introduced along the centerline of SIW rectangular cavities, which could facilitate the transition of the resonance frequency from TE_{101} mode towards TE_{201} mode, thereby forming dual-mode resonant cavities. Additionally, the adjustment on the offset distance between ports and the centerline can be used to suppress undesired spurious modes. The metallized vias are placed to perturb TE modes and to extend the stopband. The fourth-order bandpass filter can be formed by cascading two dual-mode SIW cavities. The snake-shaped slot etched on the top layer of vias wall of the adjacent SIW cavities can achieve cross-coupling and generate a pair of TZs to enhance the frequency selectivity.

The mode characteristics of SIW cavities are investigated in this paper. Three SIW filters are designed, demonstrating the design flexibility. Experimentally, a fourth-order SIW filter with a pair of TZs is fabricated and characterized. The results validate the effectiveness of the design and demonstrate its significant application potential in the field of high-performance, low-loss filters, particularly suitable for high-frequency wireless communication systems such as 5G and IoT applications. An innovative solution is provided to address the technical challenges in modern wireless communication systems.

The paper is organized as follows. Section II demonstrates the design of a second-order wide stopband SIW bandpass filter (BPF). The design of a fourth-order wide stopband SIW BPF is presented in section III. The prototype design, fabrication, and testing of a fourth-order cross-coupled SIW BPF is discussed in section IV. A specific discussion is provided in section V. Finally, a concise conclusion is provided in section VI.

II. SECOND-ORDER SIW BPF

To produce dual-mode characteristic, an isolation vias wall is added along the centerline AA' in the structure, as shown in Fig. 1 (a). A coupling window is opened on the vias wall to generate perturbation, forming dual-mode characteristic. The length of the coupling window is W_{12}.The diameter of all metal vias on the SIW sidewalls in this paper is $d = 0.4$ mm, with a spacing of $p = 0.8$ mm between adjacent vias. The coupling window placed on the central horizontal line can be used to suppress TE_{102} and TE_{202} modes. The misaligned Port1 and Port2 are used to suppress TE_{301} and TE_{401} modes in the resonator. The suppression effect is achieved by adjusting the offset S appropriately. The positions of input and

output ports of the filter are offset separately in opposite directions from the horizontal centerline (white dashed line) of the rectangular cavity, with a displacement S from the centerline of the metal vias wall. The SIW cavity is transitioned to the input and output ports using coplanar waveguides.

The resonance frequency of the TE_{m0n} mode in the proposed SIW cavity can be calculated based on:

$$f_{TE_{m0n}} = \frac{c}{2\sqrt{\mu_r \varepsilon_r}} \sqrt{\left(\frac{m}{W_1}\right)^2 + \left(\frac{n}{L_1}\right)^2}, \quad (1)$$

where c is the speed of light in a vacuum, μ_r and ε_r are the relative magnetic permeability and relative permittivity, respectively, and m and n are the mode indices along the two directions of the cavity. A Rogers RT/period 5880 substrate is used in SIW cavity design, with thickness h = 0.254 mm, relative dielectric constant $\varepsilon_r = 2.2$, and loss tangent tan$\delta = 0.0009$.

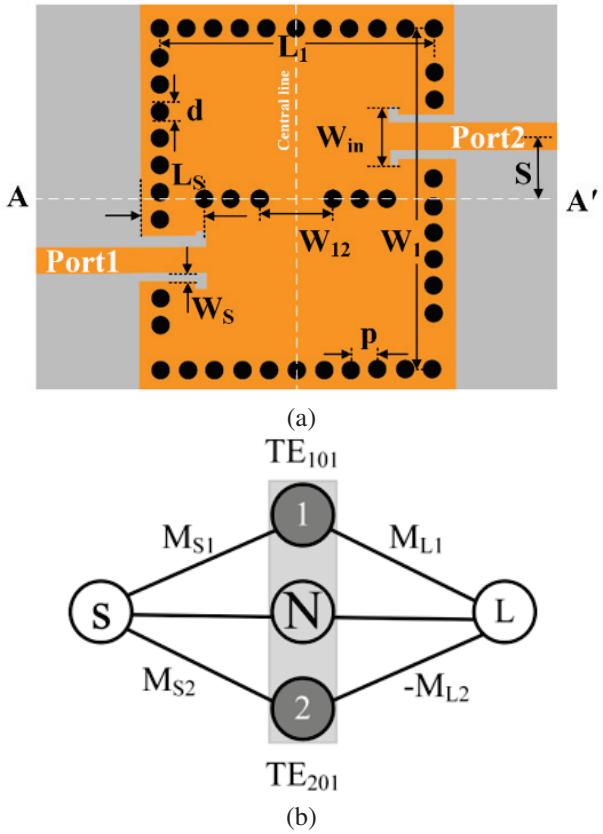


Fig. 1. (a) Geometric configuration and (b) topological structure of the dual-mode wide stopband SIW BPF.

The topology of the proposed dual-mode SIW BPF is shown in Fig. 1 (b). Nodes 1 and 2 represent TE_{101} and TE_{201} modes, respectively. The corresponding coupling matrix is shown in (2). Nodes S and L denote the input and the output ports. The electric field distributions of TE_{101} and TE_{201} modes are illustrated in Figs. 2 (a) and (b). The TE_{101} mode exhibits even symmetry in the electric field distribution, whereas the TE_{201} mode exhibits odd symmetry. Increasing the offset distance S enhances the coupling to the TE_{201} mode (M_{S2}) and reduces the coupling to the TE_{101} mode (M_{S1}) because the electric field intensity of the TE_{101} mode is the highest at the centerline, and it decays away from the centerline. Oppositely, TE_{201} mode shows zero electric field at the centerline, and is increased away from the centerline. When $M_{S1} \neq M_{S2}$, TZs can be introduced outside the passband, with their position determined by the ratio M_{S2}/M_{S1}. The ratio can be controlled by the offset distance S. Because M_{S1} is always greater than M_{S2}, and f_{TE101} is always less than f_{TE201}, TZs are positioned in the upper stopband. Additionally, node N is a non-resonant node. It is mainly used for bypass coupling created by TE_{102} mode and can generate an additional TZ in the upper stopband:

$$M = \begin{bmatrix} 0 & 0.935 & -0.935 & 0 \\ 0.935 & 1.99 & 0 & 1.064 \\ -0.935 & 0 & -1.99 & 0.935 \\ 0 & 1.064 & 0.935 & 0 \end{bmatrix}. \quad (2)$$

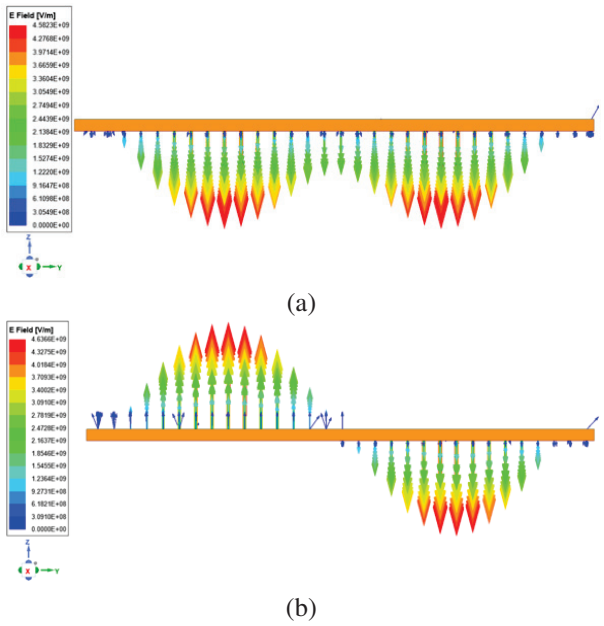


Fig. 2. Electric field distributions of (a) TE_{101} and (b) TE_{201} modes.

The coupling between the two mode cavities is mainly controlled by the width W_{12}. The coupling coefficient can be extracted by the following relation [23]:

$$k = \frac{f_1^2 - f_2^2}{f_1^2 + f_2^2}, \quad (3)$$

where f_1 and f_2 are the high and low resonant frequencies, and k is the coupling coefficient between the dual-mode cavities.

The dual-mode cavity, mainly constructed by the coupling window with width W_{12}, can produce magnetic coupling. The coupling coefficient can be extracted with the dual-mode cavity structure shown in Fig. 1 (a). Based on HFSS software, the magnetic coupling coefficient design curve is shown in Fig. 3. It can be seen that the magnetic coupling coefficient is changed with the increment of W_{12}. With W_{12} in the range of 0-1.5 mm, little change can be seen on the coupling coefficient, indicating weak coupling in this region. When W_{12} is larger than 1.5 mm, a significant rise in the coupling coefficient is observed, showing a strong increased coupling effect and a more notable interaction between the dual-mode cavities. Based on the above analysis, W_{12} is one of the key parameters for the adjustment of the magnetic coupling effect.

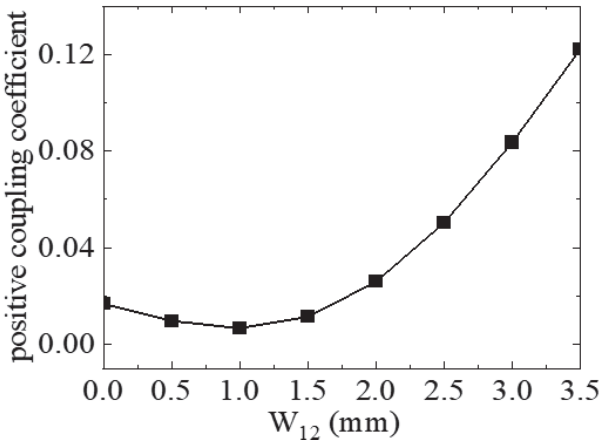


Fig. 3. Positive coupling coefficient of coupling structure.

To demonstrate complete control over the positions of TZs, an analysis of the offset distance parameter S is conducted. Figure 4 depicts the transmission characteristics S_{21} with various values of S. It can be observed that two TZs are present in the upper stopband, while no TZs are found in the lower stopband. With S increasing, TZ_1 gradually shifts towards higher frequencies with minimal variation, whereas TZ_2 shifts towards lower frequencies with more noticeable changes. As mentioned earlier, the position of the first TZ is mainly determined by the ratio M_{S2}/M_{S1}, which can be controlled by modifying the offset S. The smaller the offset S is, the smaller the M_{S2}/M_{S1} is, and the closer TZ is to the passband. The location of the second TZ depends primarily on the bypass coupling through the TE_{102} mode. As the offset S decreases, the coupling to the TE_{102} mode increases, demonstrating that S is a critical parameter for tuning TZ positions and optimizing stopband suppression.

Figure 5 shows the influences of different L_1 on the stopband of the SIW filter. It is observed that, with fixed

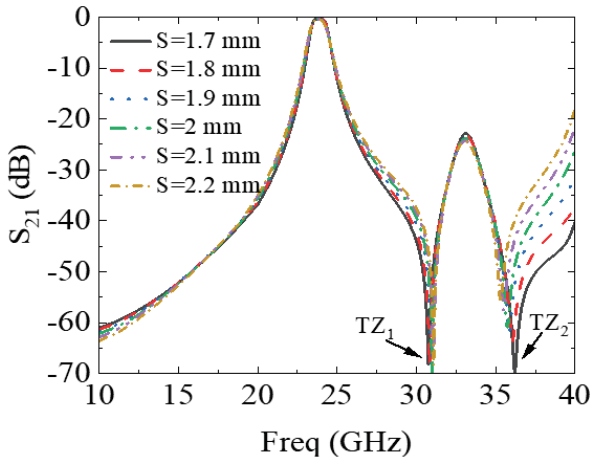


Fig. 4. Effects of different S on the TZ positions of the SIW filter.

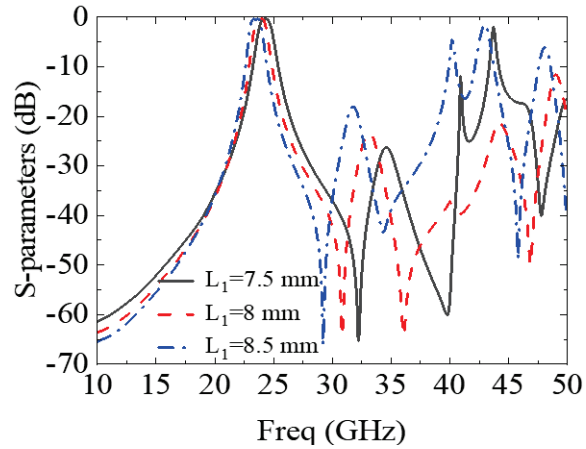


Fig. 5. Effects of different L_1 on the stopband of SIW filters.

W_1 and increased L_1, the overall frequency response is shifted towards lower frequencies. When $L_1 = 8$ mm, the suppression of f_{TE301} can be reduced to be less than 20 dB, and the coupling of TE_{301} and TE_{401} modes is weak. At 48.2 GHz, the stopband can be extended to be $2.02f_0$. When L_1 is equal to 7.5 mm or 8.5 mm, the TE_{301} and TE_{401} modes are coupled to each other, creating a parasitic passband.

The perturbed vias are used to further expand the stopband, as illustrated in Fig. 6. These perturbed vias act as metal electric walls, without electric field energy presented. Thus, the electric field distribution of transmission modes within the resonant cavity can be changed. The S-parameters of the second-order SIW BPF based on perturbation via over a wide frequency range and passband range are shown in Fig. 7. The final dimensions are: $W_1 = 10.16$ mm, $L_1 = 8$ mm, $W_{12} = 2.2$ mm, $S = 1.64$ mm, $L_S = 1.87$ mm, $W_S = 0.27$ mm,

W_{in} = 1.65 mm, Offset = 1 mm. In Fig. 7, the center frequency f_0 is 24.1 GHz, with the fractional bandwidth (FBW) of 4.7%, return loss better than 23 dB, and insertion loss below 0.36 dB. At the stopband frequency of 34 GHz, the TE_{102}/TE_{202} mode achieves suppression depth below 28.8 dB. With out-of-band suppression level below 20 dB, the stopband width reaches 53.5 GHz, corresponding to $2.22f_0$, thereby achieving significant stopband width extension.

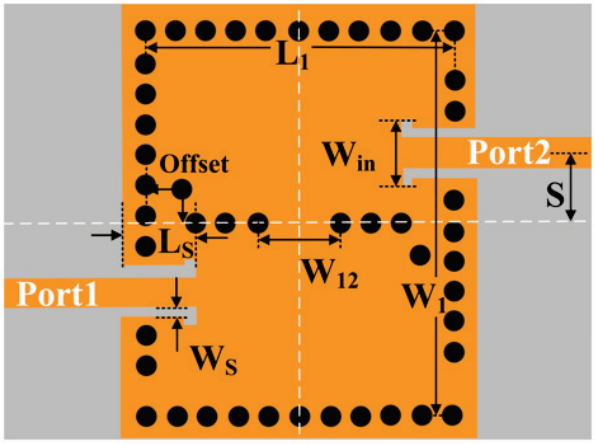


Fig. 6. Geometric configuration of second-order wide stopband SIW BPF with perturbed vias.

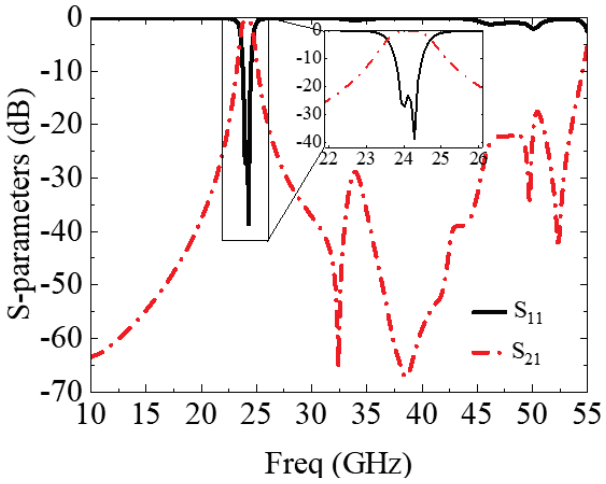


Fig. 7. Passband and wideband S-parameters of a second-order SIW BPF with perturbed vias.

III. FOURTH-ORDER SIW BPF

A fourth-order wide stopband SIW BPF is implemented by cascading two similar dual-mode SIW cavities, with the structure shown in Fig. 8 (a). The topology of the proposed fourth-order SIW BPF is shown in Fig. 8 (b). The corresponding coupling matrix is shown in (4). With the designed structure shown in Fig. 8 (a), Port1 and Port2 are symmetrically positioned relative to the centerline (white dashed line) of the vertical vias wall. Internal coupling between the two dual-mode SIW cavities is achieved through a coupling window offset from the vertical centerline of both cavities. Both left and right pairs of cavities are excited simultaneously by TE_{101} and TE_{201} modes. The coupling strength of TE_{101} and TE_{201} modes in the two cascaded structures is determined by the width W_{23} of the coupling window on the vertical vias wall. The coupling window with a width of W_{23} is arranged along the centerline between R2 and R3, which can suppress higher-order resonance modes:

$$M = \begin{bmatrix} 0 & 1.17 & 0 & 0.04 & 0 & 0 \\ 1.17 & -0.05 & 1.06 & 0 & 0.1 & 0 \\ 0 & 1.06 & -0.05 & 0.787 & 0 & 0.02 \\ 0.04 & 0 & 0.787 & -0.05 & 1.06 & 0 \\ 0 & 0.1 & 0 & 1.06 & -0.05 & 1.17 \\ 0 & 0 & 0.04 & 0 & 1.17 & 0 \end{bmatrix}. \quad (4)$$

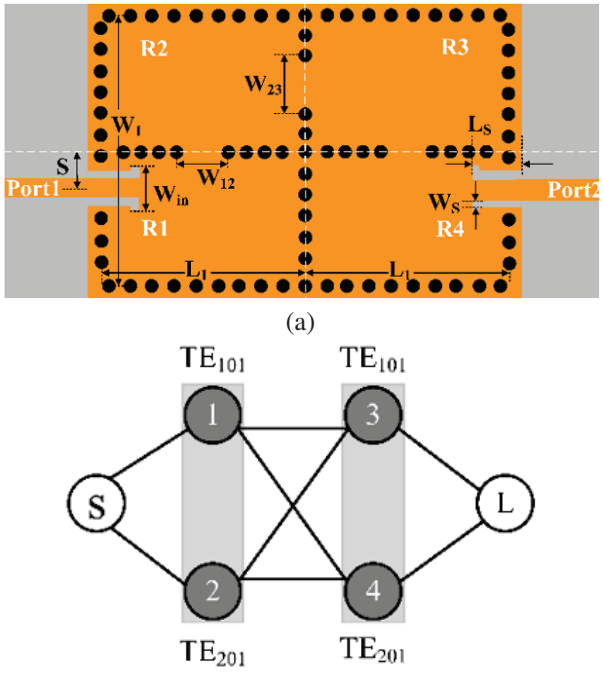


Fig. 8. (a) Geometric configuration and (b) topological structure of the fourth-order broadband SIW BPF.

As mentioned earlier, L_1 has an impact on the stopband of SIW filters. Therefore, in order to achieve a wide stopband fourth-order SIW filter, L_1 is designed to be 8 mm. Figure 9 illustrates the passband and wide stopband S-parameters of the fourth-order SIW BPF. The final dimensions are: W_1 = 10.16 mm, L_1 = 8 mm,

W_{12} = 1.8 mm, W_{23} = 2.15 mm, S = 1.4 mm, L_S = 1.88 mm, W_S = 0.27 mm, W_{in} = 1.65 mm. As depicted in Fig. 9, within the passband, frequency f_0 is 24.1 GHz, with FBW of 2.32%, return loss better than 26.85 dB, and insertion loss less than 1.26 dB. Within the stopband, the TE_{102}/TE_{202} modes exhibit attenuation greater than 26.3 dB at 33.4 GHz. With the out-of-band suppression level below 20 dB, the stopband extends to 55.65 GHz, corresponding to 2.31f_0, demonstrating effective widening of the stopband.

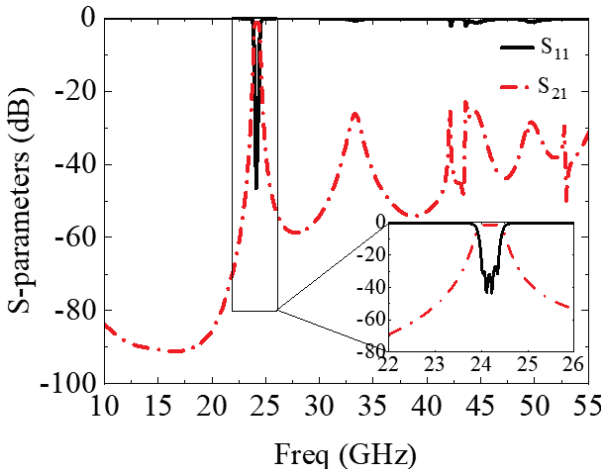


Fig. 9. Passband and wide stopband S-parameters of fourth-order SIW BPF.

IV. FOURTH-ORDER CROSS-COUPLED SIW BPF

To enhance the roll-off characteristics and improve frequency selectivity, snake-shaped slots are introduced into the filter design. These slots serve a dual purpose: they modify the coupling between adjacent resonant cavities and generate TZs near the passband edges.

In conventional SIW filters, resonant cavities are coupled positively, allowing signals to propagate in phase. This results in a relatively wide transition band and limits the filter's ability to meet the stringent selectivity demands of millimeter-wave applications. By contrast, the introduction of snake-shaped slots establishes a negative coupling structure, which alters the signal propagation path and induces phase inversion near the passband. This phase inversion leads to destructive interference with the main transmission path, effectively generating TZs in the frequency response.

The presence of these TZs significantly improves the filter's performance in two key ways. First, they increase the attenuation rate at the passband edges, effectively sharpening the transition from passband to stopband. This steep roll-off helps suppress unwanted interference from adjacent frequencies, enhancing out-of-band rejection. Second, the reduced transition bandwidth improves frequency selectivity, which is particularly crucial in millimeter-wave applications where spectrum resources are limited and efficient signal isolation is required.

Compared to designs without TZs, the inclusion of snake-shaped slots leads to a substantially steeper roll-off rate and improved frequency selectivity. This makes the filter better suited for applications demanding high spectral purity and minimal adjacent-channel interference.

Based on the fourth-order wide stopband BPF shown in Fig. 8 (a), a snake-shaped coupling slot is etched on the vias wall top layer between SIWs R1 and R4, thereby introducing negative cross-coupling. The serpentine slot design is capable of effectively introducing coupling while ensuring excellent frequency selectivity and low radiation loss without compromising the integrity and shielding effectiveness of the SIW cavity. This characteristic is especially important in high-frequency applications, as high selectivity helps to effectively separate signals from interference, ensuring signal quality and stability. In contrast, other geometries, such as rectangular and circular slots, while widely used in certain cases, often face limitations in achieving TZs, which may result in higher radiation loss or insertion loss. Furthermore, these designs may require more complex manufacturing processes, thereby affecting production efficiency. The geometric configuration and the topology are illustrated in Figs. 10 (a) and (b). Weak coupling is formed between paths 3 and 4, creating a phase difference between the two signal paths and generating a set of TZs.

The corresponding coupling matrix is shown in (5). A serpentine slot is loaded between the holes of the SIW cavity connecting the input and output ports to form electrical coupling. HFSS software extracts coupling coefficients based on the serpentine groove structure. The electrical coupling coefficient design curve is shown in Fig. 11. It can be seen that the negative coupling coefficient is decreased with W_e increasing, indicating that the negative coupling strength is gradually increased. As W_e is increased from 0.25 mm to 0.65 mm, the negative coupling effect between the two cavities is significantly enhanced, and the coupling strength becomes stronger:

$$M = \begin{bmatrix} 0 & 1.17 & 0 & 0.02 & 0 & 0 \\ 1.17 & -0.05 & 1.07 & 0 & 0.08 & 0 \\ 0 & 1.07 & -0.05 & 0.82 & 0 & 0.02 \\ 0.02 & 0 & 0.82 & -0.05 & -1.06 & 0 \\ 0 & 0.08 & 0 & -1.06 & -0.05 & 1.16 \\ 0 & 0 & 0.02 & 0 & 1.16 & 0 \end{bmatrix}. \quad (5)$$

The final dimensions of fourth-order cross-coupled SIW BPF are: W_1 = 10.2 mm, L_1 = 7.6 mm,

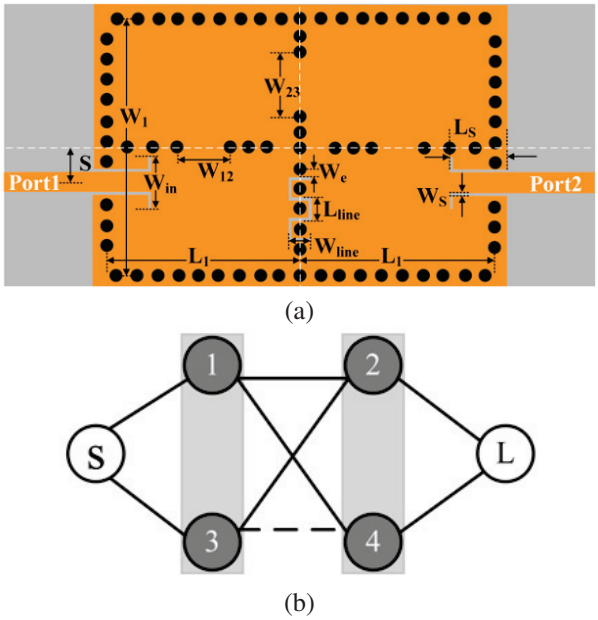


Fig. 10. (a) Geometric configuration and (b) topological structure of the fourth-order cross-coupled SIW BPF.

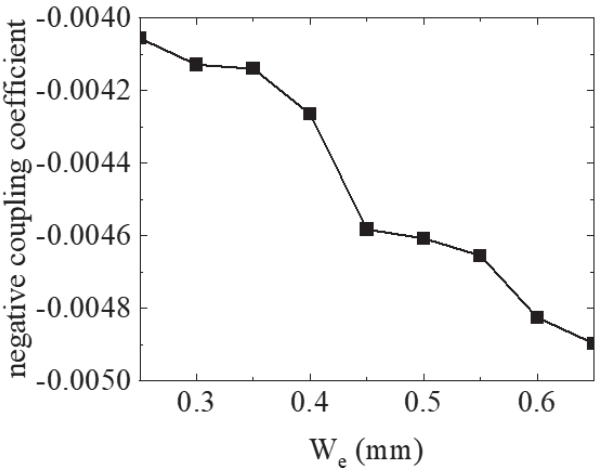


Fig. 11. Negative coupling coefficient of coupling structure.

$W_{12} = 2.1$ mm, $W_{23} = 2.6$ mm, $S = 1.5$ mm, $L_S = 2.24$ mm, $W_S = 0.1$ mm, $W_{in} = 2.05$ mm, $W_e = 0.35$ mm, $W_{line} = 0.85$ mm, $L_{line} = 0.9$ mm. The structure diagram of the prototype fourth-order cross-coupled SIW BPF is shown in Fig. 12. Figure 13 illustrates simulated and measured responses. The measured results are in good agreement with the simulation results. The simulated center frequency f_0 is 24.1 GHz with a minimum insertion loss of 0.75 dB and return loss better than 20.18 dB. Two TZs are observed at 22.7 GHz and 25.14 GHz, with suppressions of 54.65 dB and 37.95 dB, respectively. The measured center frequency f_0 is 24.1 GHz

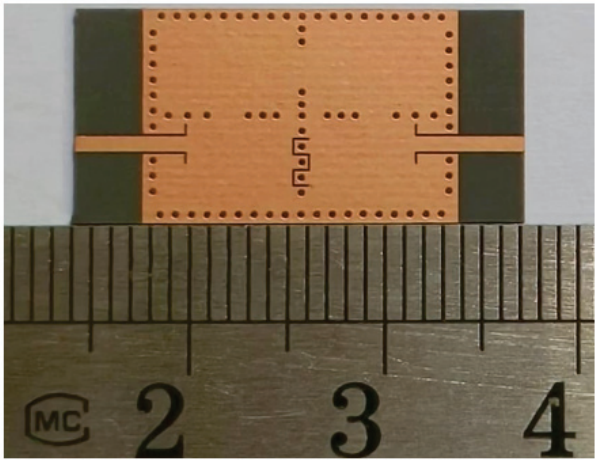


Fig. 12. Photograph of the fourth-order cross-coupled SIW BPF.

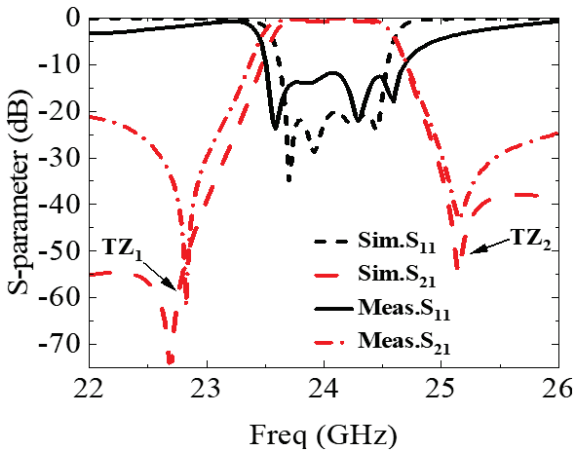


Fig. 13. S-parameters of the fourth-order cross-coupled SIW BPF.

with a minimum insertion loss of 0.52 dB and return loss better than 11.8 dB, revealing TZs at 22.82 GHz and 25.15 GHz with suppression levels below 20 dB.

The out-of-band suppression is better than 40 dB in the simulation and better than 20 dB in the actual measurement. Although the measured out-of-band suppression is lower than the simulated result, the TZ in the lower stopband is closer to the passband, showing better frequency selectivity than the simulation. The insertion loss is further decreased after loading the coupling slot. The difference of the return loss between the simulation and the measurement in Fig. 13 can be attributed to manufacturing errors, soldering of SMA connectors, and parasitic effects from test fixtures, connectors, and measurement equipment. These factors can introduce additional losses and reflections, which reduce the out-of-band suppression effectiveness. During the simulation process, certain physical phenomena may not have been

fully included, such as the parasitic effects at high frequencies and the non-ideal characteristics of the substrate materials. Additionally, the materials used in actual measurements may differ from the simulations, such as variations in dielectric constant and loss factor. The testing environment may also impact the results. Changes in temperature and humidity, as well as the calibration status of the measurement equipment, may lead to errors in the results. Furthermore, the connection methods and layout design of the test fixtures may also affect signal transmission and measurement to some extent:

$$Q_u = \frac{f_0}{BW} \times Q_{ext}. \quad (6)$$

According to (6), when f_0 is 24.1 GHz, the bandwidth is 1.133 GHz, Q_{ext} is 21.26, and Q_u is 452.

After TZs through serpentine slots to optimize the roll-off characteristics at the passband edges, attention should still be given to the suppression of higher-order frequencies. Although TZs effectively mitigate interference near the passband, their influence is primarily concentrated at specific frequency points, with limited impact on higher-order resonant frequencies. In millimeter-wave SIW filters, these higher-order resonant frequencies can introduce unwanted interference outside the passband, which may degrade overall system performance, particularly in spectrally dense environments. Therefore, the suppression of higher-order frequencies is regarded as an important step in enhancing the overall performance of the filter.

To more effectively attenuate higher-order frequencies, the design of the serpentine slots can be optimized. Specifically, as shown in Fig. 14, reducing the diameter of the metal vias above and below the serpentine slots, and increasing the distance between the vias and the slots, can significantly reduce the coupling effect of the serpentine slots on higher-order modes. This design adjustment effectively mitigates the excitation of higher-order modes, further optimizing the filter's high-frequency attenuation characteristics.

The diameter and positioning of the metal vias around the serpentine slots are shown in Fig. 15, which presents four distinct configurations. The first configuration represents the initial case, where the via diameter is $d_1 = d = 0.4$ mm, and the vias are positioned at the standard distance from the serpentine slot, with $W_{e1} = W_{e2} = 0.35$ mm. The second configuration reduces the via diameter to $d_1 = 0.3$ mm, while maintaining $W_{e1} = W_{e2} = 0.35$ mm. The third configuration increases the distance between the vias and the serpentine slot, with $W_{e1} = 0.53$ mm, $W_{e2} = 0.84$ mm, and the via diameter remaining at $d_1 = 0.4$ mm. The final configuration combines both optimizations, where the vias are moved further away from the serpentine slot ($W_{e1} = 0.53$ mm, $W_{e2} = 0.84$ mm), and the via diameter is reduced to $d_1 = 0.3$ mm.

By comparing these four configurations, it can be observed that both reducing the via diameter and increasing the distance between the vias and the serpentine slot effectively mitigate the impact of higher-order modes. Specifically, reducing the via diameter helps weaken the coupling between higher-order modes, while distancing the vias from the serpentine slot reduces unnecessary interference signals. When these two optimization strategies are combined, they significantly suppress the influence of higher-order frequencies, enhancing the filter's ability to reject interference and out-of-band signals at higher frequencies, thereby improving the overall performance of the filter.

However, it is important to note that this optimization leads to a decrease in the filter's selectivity. In other words, although the suppression of higher-order frequencies is enhanced, the corresponding selectivity worsens. This trade-off is an inevitable consequence of the attenuation process of higher-order frequencies.

Figure 16 shows the passband and wide stopband S-parameters of a fourth-order SIW BPF with suppressed higher-order modes. The final dimensions are: $W_1 = 10.16$ mm, $L_1 = 7.5$ mm, $W_{12} = 2.19$ mm, $W_{23} = 2.61$

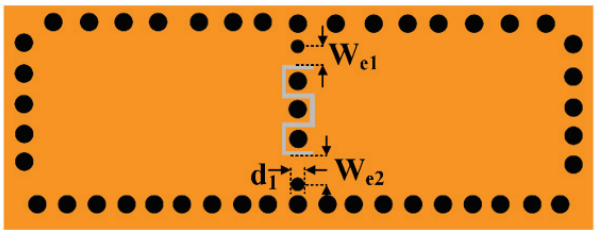


Fig. 14. Optimization design of the serpentine slot section.

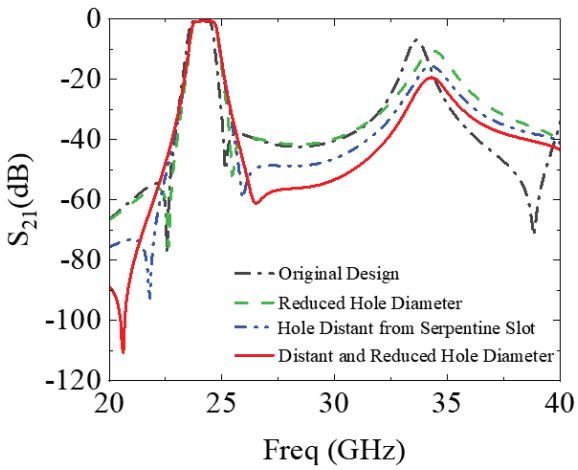


Fig. 15. Effect of via hole diameter and position on serpentine slot performance.

mm, S = 1.99 mm, L_S = 2.27 mm, W_S = 0.1 mm, W_{in} = 1.69 mm, W_{e1} = 0.53 mm, W_{e2} = 0.84 mm, d_1 = 0.3 mm, W_{line} = 1 mm, L_{line} = 0.9 mm. As shown in Fig. 16, within the passband, the frequency f_0 is 24.1 GHz, with FBW of 4.5%, return loss better than 20 dB, and insertion loss less than 0.7 dB. In the stopband, the attenuation of TE_{102}/TE_{202} modes exceeds 20 dB at 34.2 GHz. The stopband extends to 40 GHz ($1.66f_0$) when the out-of-band suppression level is below 20 dB, indicating an effective widening of the stopband.

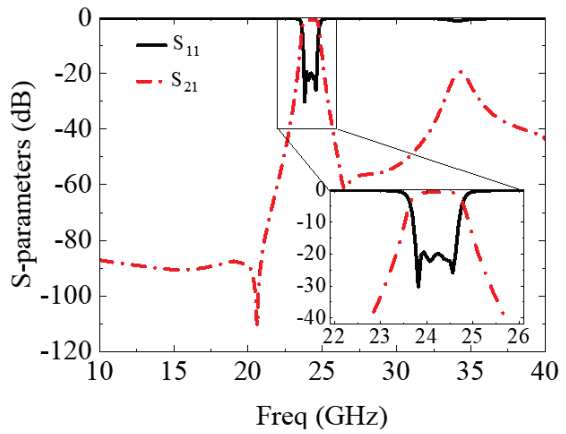


Fig. 16. The passband and wide stopband S-parameters of high-order suppressed fourth-order SIW BPF.

Comparison between the proposed fourth-order cross-coupled SIW BPF and the published results is listed in Table 1. It can be seen that the design in this paper shows more compact size and lower insertion loss (IL). A filter with narrower bandwidth can provide better frequency selectivity, allowing the filter to effectively distinguish the target signal from interference. This is particularly important in signal processing, especially in high-frequency communication systems such as millimeter-wave communication and radar applications, which typically have high demands for frequency selectivity and insertion loss. Additionally, a moderate bandwidth design is beneficial in reducing insertion loss, ensuring effective signal transmission and overall system performance. Therefore, although modern communication systems generally pursue wider bandwidths, a careful balance between the signal quality and the transmission stability is considered in our design at specific frequencies.

It is noteworthy that the size of the design in [23] is smaller than our design. However, its insertion loss is 0.9 dB, with narrower bandwidth. In [24–27], those designs are limited to lower frequency bands with higher ILs. Additionally, the miniaturization achieved in [28] is due to a complex multilayer structure, whereas the planar layout in the paper is simpler. The inflatable SIW filter used in [29] can achieve a wide bandwidth but it has a very high insertion loss of 4 dB and does not include TZs. Additionally, the manufacturing process for inflatable SIW filters is often complex, requiring strict control over materials and structures, which may lead to higher production costs. Furthermore, the stability of inflatable SIW filters during long-term use also needs to be considered, particularly in high-frequency applications where even minor structural changes can affect the performance. In contrast, the design proposed in this paper, based on a single-layer PCB structure, is easy to fabricate, allowing for the realization of a high-performance filter with low cost. Moreover, the stability of this design in high-frequency signal transmission is effectively ensured.

V. DISCUSSION

In SIW filter design, electrical coupling is achieved by using metal surface slits, which is a common and effective technique. The surface current of the TE_{101} mode is significantly suppressed by the open slot-line structures. As a result, radiation losses are introduced, with reduced shielding effectiveness and degraded quality factors.

Table 1: Comparison of this design with other types of filters

Ref.	f_0 (GHz)	FBW (%)	Size (λ_g^2)	IL (dB)	NP	NZ	Layers	Size (mm^2)
[24]	20.5	4	1.93	0.9	4	2	1	14×13.43
[25]	5	2.9	1.30	6.35	4	2	1	35.9×35.7
[26]	7.5	5.2	1.15	1.5	3	2	1	52.4×16
[27]	10	3.3	2.97	1.55	4	2	1	39.4×30.9
[28]	9.95	3.1	2.09	2.1	4	2	1	30×28.81
[29]	26.8	7	1.74	2.1	4	4	2	10.9×8.52
[30]	5.45	11.2	1.73	4	2	0	1	96.5×49.4
This work	24.1	4.7	2.20	0.52	4	2	1	15.2×10.2

f_0: center frequency; FBW: fractional bandwidth; IL: insertion loss; NP: number of poles; NZ: number of TZs.

Table 2: Comparison of this design with other cross-coupled structures

Ref.	f_0 (GHz)	FBW (%)	Size (λ_g^2)	IL (dB)	NP	NZ	Layers	Size (mm^2)
[31]	2.2	5.5	0.47	1.55	4	2	4	49.2×46
[32]	26.1	10.5	8.91	2.65	6	2	1	22×21
[33]	12.5	9	7.07	1.3	6	2	1	23.5×17.5
[34]	5.75	2	2.12	2.2	4	2	1	36×35.7
[35]	4.5	3.7	0.51	2.1	4	3	2	25×25
[36]	38.8	5.9	4.38	6.7	6	2	2	13.1×5.75
This work	24.1	4.7	2.20	0.52	4	2	1	15.2×10.2

f_0: center frequency; FBW: fractional bandwidth; IL: insertion loss; NP: number of poles; NZ: number of TZs.

Table 2 shows the comparison between the fourth-order cross-coupled SIW BPF proposed in this paper and other reported cross-coupled designs. Significant advantages are shown by the filter design in this paper, and it is seen to be more competitive compared to the designs in other literature. Serpentine slots are introduced on copper pillar walls, which do not cause large radiation losses. An insertion loss of only 0.52 dB is achieved, which is much lower than the 2.65 dB reported in [31] and the 2.2 dB in [34], indicating excellent signal transmission efficiency. Although the physical size is 15.2×10.2 mm^2, a wavelength area of $2.20\lambda_g^2$ is reached. This is more compact and of greater efficiency compared to other designs with larger size, e.g. $8.91\lambda_g^2$ in [31] and $7.07\lambda_g^2$ in [32]. Additionally, a single-layer structure is used in this paper, which greatly simplifies the manufacturing process. Compared to the four-layer structure used in [30], both the complexity and the cost are reduced, and easier integration into various systems is achieved. The filter in this paper shows advantage over the design in [33]. Both the designs in this paper and in [33] show similar structure, except a low-frequency band with a narrower bandwidth is applied in [33].

Specifically, two significant transmission zeros (TZ_1 and TZ_2) are successfully introduced in the filter design of this paper, located near 22.5 GHz and 25 GHz. These TZs greatly improve the filter's frequency selectivity. At the TZ locations, a suppression depth of over 60 dB is achieved, which is much better than that of the multi-layer cross-coupled design shown in [35]. Although some out-of-band suppression capability is shown by the multi-layer cross-coupled filters, no clear TZs are exhibited. No TZs are seen in the lower stopband, and the upper stopband has a TZ that is not very noticeable. The simulation results of the filter in this paper coincide with the measured data closely across the entire frequency range, especially near the TZs, indicating high consistency and reliability in design and manufacturing. In contrast, while basic agreement between measurements and simulations is shown by the multi-layer filters, noticeable deviations are found in certain frequency ranges (such as around 37 GHz), particularly in out-of-band suppression.

Additionally, better performance in insertion loss is shown by the filter in this paper, which maintains low insertion loss and improves overall signal transmission quality. In contrast, higher insertion loss is found in the multi-layer filter design, and out-of-band suppression has room for improvement. Thus, clear advantages in frequency selectivity, suppression performance, and consistency are demonstrated by the filter design in this paper, making it more suitable for high-demand communication systems.

VI. CONCLUSION

In this paper, a novel fourth-order bandpass filter based on a cross-coupled structure is proposed. The filter shows a pair of TZs, thereby enhancing frequency selectivity. The resonant characteristics of the first six resonant modes in the proposed single-multi-mode SIW cavity is studied, validating the feasibility of a bandpass filter in terms of frequency and bandwidth control. Based on this exploration, three SIW filters are designed and studied, demonstrating the wide stopband characteristics of the second-order and the fourth-order SIW bandpass filters. Furthermore, a fourth-order SIW filter with a pair of TZs is designed, fabricated, and measured, showing good agreement between the simulated and the measured results. This technology stands as a competitive candidate for developing high-performance SIW filters in microwave/radio circuit and system applications, particularly suitable for high-frequency wireless communication systems such as 5G and IoT applications.

ACKNOWLEDGMENT

This research was funded by the Fujian Natural Science Foundation Project grant number 2022J01823 and Technology Subsidy Project (2024CXY0314).

REFERENCES

[1] P. Li, H. Chu, and R.-S. Chen, "Design of compact bandpass filters using quarter-mode and eighth-mode SIW cavities," *IEEE Trans. Compon. Packag.*

Manuf. Technol., vol. 7, no. 6, pp. 956-963, June 2017.

[2] Z. Liu, G. Xiao, and L. Zhu, "Triple-mode bandpass filters on CSRR-loaded substrate integrated waveguide cavities," *IEEE Trans. Compon. Packag. Manuf. Technol.,* vol. 6, no. 7, pp. 1099-1105, July 2016.

[3] X. Zhou, G. Zhang, S. Feng, K. W. Tam, Z. Zhang, and W. Tang, "Design of 3-D integrated SIW multiband bandpass filter with split-type extended doublet topology," *IEEE Trans. Compon. Packag. Manuf. Technol.,* vol. 12, no. 10, pp. 1681-1691, Oct. 2022.

[4] X.-P. Chen and K. Wu, "Substrate integrated waveguide filter: Basic design rules and fundamental structure features," *IEEE Microw. Mag.,* vol. 15, no. 5, pp. 108-116, July/Aug. 2014.

[5] Y. Zheng, Y. Zhu, Z. Wang, and Y. Dong, "Compact, wide stopband, shielded hybrid filter based on quarter-mode substrate integrated waveguide and microstrip line resonators," *IEEE Microw. Wireless Compon. Lett.,* vol. 31, no. 3, pp. 245-248, Mar. 2021.

[6] P. Chu, L. Guo, L. Zhang, F. Xu, W. Hong, and K. Wu, "Wide stopband substrate integrated waveguide filter implemented by orthogonal ports' offset," *IEEE Trans. Microw. Theory Techn.,* vol. 68, no. 3, pp. 964-970, Mar. 2020.

[7] G. Lin and Y. Dong, "A compact, hybrid SIW filter with controllable transmission zeros and high selectivity," *IEEE Trans. Circuits Syst. II Exp. Briefs,* vol. 69, no. 4, pp. 2051-2055, Apr. 2022.

[8] Y. Zhu and Y. Dong, "A compact dual-band quasi-elliptic filter based on hybrid SIW and microstrip technologies," *IEEE Trans. Circuits Syst. II Exp. Briefs,* vol. 69, no. 3, pp. 719-723, Mar. 2022.

[9] W. Shen, "Extended-doublet half-mode substrate integrated waveguide bandpass filter with wide stopband," *IEEE Microw. Wireless Compon. Lett.,* vol. 28, no. 4, pp. 305-307, Apr. 2018.

[10] Y. Zhu and Y. Dong, "A compact dual-band quasi-elliptic filter based on hybrid SIW and microstrip technologies," *IEEE Trans. Circuits Syst. II Exp. Briefs,* vol. 69, no. 3, pp. 719-723, Mar. 2022.

[11] G. Lin and Y. Dong, "A compact, hybrid SIW filter with controllable transmission zeros and high selectivity," *IEEE Trans. Circuits Syst. II Exp. Briefs,* vol. 69, no. 4, pp. 2051-2055, Apr. 2022.

[12] L. Gu and Y. Dong, "Compact half-mode SIW filter with high selectivity and improved stopband performance," *IEEE Microw. Wireless Compon. Lett.,* vol. 32, no. 9, pp. 1039-1042, Sep. 2022.

[13] M. R. Jiao, F. Zhu, P. Chu, W. Yu, and G. Q. Luo, "Compact hybrid bandpass filters using substrate-integrated waveguide and stripline resonators," *IEEE Trans. Microw. Theory Techn.,* vol. 72, no. 1, pp. 391-400, Jan. 2024.

[14] Y.-X. Yan, W. Yu, and J.-X. Chen, "Millimeter-wave low side- and back-lobe SIW filtenna array fed by novel filtering power divider using hybrid TE_{101}/TE_{301} mode SIW cavities," *IEEE Access,* vol. 9, pp. 167706-167714, Dec. 2021.

[15] G. Lin, Y. Dong, and X. Luo, "Miniaturized quarter-mode SIW filters loaded by dual-mode microstrip resonator with high selectivity and flexible response," *IEEE Microw. Wireless Compon. Lett.,* vol. 32, no. 6, pp. 660-663, June 2022.

[16] Z. Chen, M. Wang, D. Guan, Z. Qian, W. Wu, and L. Zhu, "Wideband filtering antenna fed through hybrid substrate integrated waveguide and spoof localized surface plasmon structure," *IEEE Trans. Antennas Propa.,* vol. 70, no. 5, pp. 3812-3817, May 2022.

[17] D. Li, W. Luo, X. Chen, Y. Liu, K.-D. Xu, and Q. Chen, "Miniaturized dual-/tri-/quad-band bandpass filters using perturbed multimode SIW cavity," *IEEE Trans. Compon. Packag. Manuf. Technol.,* vol. 13, no. 10, pp. 1685-1693, Oct. 2023.

[18] M.-H. Ho and K.-H. Tang, "Miniaturized SIW cavity tri-band filter design," *IEEE Microw. Wireless Compon. Lett.,* vol. 30, no. 6, pp. 589-592, June 2020.

[19] H.-W. Xie, K. Zhou, C.-X. Zhou, and W. Wu, "Substrate-integrated waveguide triple-band bandpass filters using triple-mode cavities," *IEEE Trans. Microw. Theory Techn.,* vol. 66, no. 6, pp. 2967-2977, June 2018.

[20] P. Liu, Z. Li, M. Qin, J. Yin, and X. Qiu, "Two compact bandpass filters with controllable band based on eighth-mode substrate integrated waveguide," *IEEE Trans. Circuits Syst. II Exp. Briefs,* vol. 71, no. 2, pp. 932-936, Feb. 2024.

[21] Q. Liu and L. Zhu, "Design of cross-coupled bandpass filters with flexible coupling via half-mode substrate-integrated waveguide," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 2024, pp. 1-13, 2024.

[22] P. Chu, M. Luo, J. Zhou, L. Guo, F. Zhu, and L. Zhang, "Dual-band substrate integrated waveguide filter with independent TE_{101} and TE_{102} coupling," *IEEE Trans. Microw. Theory Techn.,* vol. 72, no. 3, pp. 1877-1885, Mar. 2024.

[23] J.-S. Hong and M. J. Lancaster, *Microstrip Filter for RF/Microwave Applications*. New York: Wiley, 2001.

[24] X.-P. Chen and K. Wu, "Substrate integrated waveguide cross-coupled filter with negative coupling structure," *IEEE Trans. Microw. Theory Techn.,* vol. 56, no. 1, pp. 142-149, Jan. 2008.

[25] C. J. You, Z. N. Chen, X. W. Zhu, and K. Gong, "Single-layered SIW post-loaded electric coupling-enhanced structure and its filter applications," *IEEE Trans. Microw. Theory Techn.,* vol. 61, no. 1, pp. 125-130, Jan. 2013.

[26] P. Chu, P. Zhu, J. Feng, L. Guo, L. Zhang, and F. Zhu, "Substrate integrated waveguide filter with flexible mixed coupling," *IEEE Trans. Microw. Theory Techn.,* vol. 71, no. 9, pp. 4003-4011, Sep. 2023.

[27] F. Zhu, G. Q. Luo, B. You, X. H. Zhang, and K. Wu, "Planar dual-mode bandpass filters using perturbed substrate-integrated waveguide rectangular cavities," *IEEE Trans. Microw. Theory Techn.,* vol. 69, no. 6, pp. 3048-3057, June 2021.

[28] A. R. Azad and A. Mohan, "Substrate integrated waveguide cross-coupled bandpass filter with wide-stopband," in *2020 URSI Regional Conference on Radio Science (URSI-RCRS),* Varanasi, India, pp. 1-4, 2020.

[29] K. Gong, W. Hong, Y. Zhang, P. Chen, and C. J. You, "Substrate integrated waveguide quasi-elliptic filters with controllable electric and magnetic mixed coupling," *IEEE Trans. Microw. Theory Techn.,* vol. 60, no. 10, pp. 3071-3078, Oct. 2012.

[30] A. Moznebi, K. Afrooz, and M. Danaeian, "High-performance filtering power divider based on air-filled substrate integrated waveguide technology," *ETRI J.*, vol. 45, no. 2, pp. 338-345, July 2022.

[31] A. A. Khan, A. Gupta, and M. K. Mandal, "Multilayer cross-coupled SIW filter with next harmonic suppression," in *2022 IEEE Microwaves, Antennas, and Propagation Conference (MAPCON)*, Bangalore, India, pp. 68-73, 2022.

[32] X. Zhang, Y. Ji, W. Chen, J. Zhou, and M. Wang, "Design of a high outband rejection cross-coupled SIW filter for millimeter wave communications," in *2019 International Conference on Microwave and Millimeter Wave Technology (ICMMT),* Guangzhou, China, pp. 1-3, 2019.

[33] D. L. Diedhiou, E. Rius, J.-F. Favennec, and A. El Mostrah, "Ku-band cross-coupled ceramic SIW filter using a novel electric cross-coupling," *IEEE Microw. Wireless Compon. Lett.,* vol. 25, no. 2, pp. 109-111, Feb. 2015.

[34] S. Sirci, M. Á. Sánchez-Soriano, J. D. Martı´nez, V. E. Boria, F. Gentili, W. Bösch, and R. Sorrentino, "Design and multiphysics analysis of direct and cross-coupled SIW combline filters using electric and magnetic couplings," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 12, pp. 4341-4354, Dec. 2015.

[35] B. Lee, G. Lee, and J. Lee, "Two-layered cross-coupled post-loaded SIW filter with microstrip ports," *IEEE Trans. Circuits Syst. II Exp. Briefs,* vol. 70, no. 4, pp. 1346-1350, Apr. 2023.

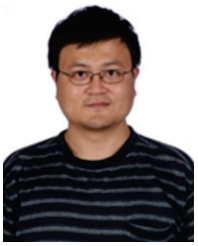
[36] G. Lee, B. Lee, J.-Y. Jeong, and J. Lee, "Ka-band surface-mount cross-coupled SIW filter with multi-layered microstrip-to-GCPW transition," *IEEE Access*, vol. 7, pp. 66453-66462, 2019.

Yawen He was born in Nanjing, China, in 2001. She received the bachelor's degree in communication engineering from Wuyi University, Nanping, China, in 2023. She is currently working toward the Postgraduate degree in Information and Communication Engineering, Jimei University, Xiamen, China. Her research interest is RF microwave millimeter wave circuit design.

Zhonghua Ma was born in Gansu, China, in 1973. He received his Ph.D. in Microelectronics from Lanzhou University, Lanzhou, China, in 2018. His present research interests include antenna techniques, RF and microwave circuits design and IoT.

Mengnan Wang was born in Pingdingshan, China, in 2000. She received the bachelor's degree in communication engineering from Zhengzhou University of Aeronautics, Zhengzhou, China, in 2022. She is currently working toward the Postgraduate degree in Information and Communication Engineering, Jimei University, Xiamen, China. Her research interest is chipless RFID sensors.

Jingyao Huang was born in Xiamen, China, in 2004. She is currently pursuing her B.S. degree with the Department of Communication Engineering, Jimei University, Xiamen, China. Her research interests are in microwave technology and rectenna technology.



Yanfeng Jiang was born in Jilin, China, in 1972. He received the B.S. degree in Electrical Engineering from Southeast University, Nanjing, China, in 1993, and Ph.D. degree in Microelectronics from Lanzhou University, Lanzhou, China, in 2000. He is now full professor in Jiangnan University, Wuxi, China. His research interests include integrated circuit design, power semiconductor devices and magnetic devices.

SIW Cavity-backed Gain-enhanced Circularly Polarized Metamaterial-loaded Dual-band MIMO Antenna for WLAN and 5G Applications

Infant Leo S[1], G. Aloy Anuja Mary[2], A. Syed Mazhar[1], Satyam Mishra[1], and G. Jothi[3]

[1]Department of Electrical, Electronics and Communication Engineering, Galgotias University
Gautam Buddh Nagar, Greater Noida, Uttar Pradesh, India
infant.s@galgotiasuniversity.edu.in, syed.mazhar@galgotiasuniversity.edu.in,
mishra.satyam@galgotiasuniversity.edu.in

[2]Department of Electronics and Communication Engineering, School of Electrical and Communication, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Avadi, Chennai-600062, India
draloyanujamary@veltech.edu.in

[3]Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences
Chennai, Tamilnadu, India
jothigopal1006@gmail.com

Abstract – This work proposes the development of a metamaterial-loaded circularly dual-band cavity-backed substrate integrated waveguide (SIW) MIMO antenna designed for the sub-6 GHz, emphasizing sub-6 GHz 5G and WLAN applications. The creation of dual operating bands is enabled via a modified dual split ring resonator (CSRR)-shaped slot that is etched into the SIW cavity-backed rectangular radiator. Additionally, the antenna incorporates 6x3 modified SSRR unit cells strategically located in front of the intended radiators along the y-axis. This arrangement enables circular polarization and enhances the gain of the proposed radiator at 3.3 GHz and 5 GHz. The metamaterial loading of the proposed antenna yields a gain of 5.5 dB at 2.4 GHz and 5.4 dB at 5 GHz. Further, the implementation of a CSRR electromagnetic bandgap (EBG) decoupling structure reduces the mutual coupling between the radiators. The antenna exhibits an exceptional diversity performance. The experimental validation of the system confirms its intended functionality.

Index Terms – 5G sub-6 GHz, circularly polarized, gain enhancement, metasurface, MIMO antenna, SIW cavity.

I. INTRODUCTION

There is a notable need for improved data transmission capacity and link reliability in the current 5G age of wireless communication technology, as wireless devices are progressing at a rapid pace [1]. Among the several methods proposed by researchers to meet demand, MIMO technology has emerged as a highly effective means of improving data rates [2]. Without utilizing extra resources, MIMO antennas (multiple-input multiple-output) can increase channel capacity and link dependability [2]. Circularly polarized antennas, in contrast, have small propagation losses because they efficiently combat multipath fading [3]. This is why modern wireless networks can't function without circularly polarized MIMO antennas. Although wideband MIMO antennas with circular polarization have been created, modern communication systems are better served by multi-band antennas with circular polarization [3]. The literature reports a plethora of methods for achieving circular polarization. Circular polarization can be achieved by strategically placing a patch on top of differently shaped slots, as described in [4, 5]. In [6, 7], the circular polarization is achieved by means of a metamaterial superstate. Many design considerations also need to be satisfied in order to achieve compact dimensions, excellent gain, and minimized mutual coupling.The literature has been updated with a variety of MIMO antennas that have been introduced by researchers in recent years [4, 6, 7]. The literature states that in order to achieve a reduction in mutual coupling, it is necessary to place an electromagnetic bandgap (EBG), parasitic elements, a rectangular stub, a various shape decoupling structure, and metamaterial [8, 9]. Over the past decade, a novel technology called substrate integrated waveguide (SIW) cavities has emerged, offering an extensive range of applications and innovative strategies for several existing applications [10, 11]. The SIW cavity-backing architecture, positioned near the edge of the ground plane structure, has significantly reduced back lobe radiation and unwanted surface currents. Thus, it enhances

Submitted On: December 11, 2024
Accepted On: March 14, 2025

https://doi.org/10.13052/2025.ACES.J.400410

antenna gain while maintaining a compact form factor and providing excellent cost-effectiveness. Conventional SIW cavity-backed slot antennas exhibit significant challenges due to their limited operating bandwidth and substantial size. Researchers have explored many ways over several years to enhance the bandwidth of the SIW cavity-backed slot antenna. The enhancement in bandwidth is achieved by exciting two closely placed mode combinations, with adjustments made to the cavity modes. This facilitates an enhancement of the bandwidth. Tuning of the cavity's position, dimensions, and slot size is necessary to optimize the cavity modes [12]. The implementation of shorting vias [12] and the slot ring [13] aims to enhance the bandwidth of the SIW antenna.

In environments with significant dispersion, such as industries, a high-gain antenna is essential. The antenna's gain is improved by the utilization of multi-layer substrates [14] and metamaterials [15] as reported in the literature. A challenge in enhancing gain is the increasing desire for smaller gadgets that occupy less physical space. Metamaterials have recently emerged as a feasible and economical approach to enhance gain without significantly increasing the physical dimensions of the antenna. Metamaterials are a unique class of synthetic compounds that exhibit extraordinary properties that are not found in organic substances. The integration of metamaterials into various antennas substantially enhances antenna gain performance [16]. The metamaterial array is situated in front of the antenna to boost the gain [16, 17]. Furthermore, the antenna gain is improved by placing the metamaterial array above the antenna [18]. The antennas that have been listed above illustrate a variety of methods for circularly polarized MIMO antenna implementation, mutual coupling reduction, bandwidth enhancement, and gain enhancement; however, there are only a few sub-6 GHz MIMO antennas that are now available that provide a wider bandwidth with a higher gain in a compact-sized planar configuration. Further, the mentioned antennas used a metamaterial layer superstate for gain improvement, but this method makes antennas bulky. In this work, the development of a metamaterial-loaded circularly dual-band cavity-backed SIW MIMO antenna designed for the sub-6 GHz frequency is emphasized, focusing on sub-6 GHz 5G and WLAN applications.

II. MIMO ANTENNA CONFIGURATION AND ITS DEVELOPMENT

A. SIW cavity-backed radiator development

Figure 1 (a) illustrates the arrangement of the proposed SIW cavity-backed slot antenna (Antenna-1). The design incorporates a rectangular SIW cavity, formed by arranging metal vias in a rectangular pattern, as illustrated in Fig. 1 (a). The traditional rectangular SIW cavity-backed antenna is designed with metallic vias to prevent electromagnetic energy leakage. The SIW standards require that the pitch distance (s) and diameter (d) of metallic vias maintain the ratios $d/s \geq 0.5$ and $d/\lambda_0 \leq 0.1$. According to the information presented in [10], the dimensions of the rectangular cavity resonator have been estimated using Equation (1), leading to a first-order mode (i.e., TE110) resonance at approximately 3.4 GHz, as illustrated in Fig. 1 (f):

$$f_r\,(TE_{mnp}) = \frac{c}{2\sqrt{\varepsilon_r}}\sqrt{\left(\frac{m}{L}\right)^2+\left(\frac{n}{W}\right)^2+\left(\frac{p}{h}\right)^2}, \quad (1)$$

where L – SIW cavity length, W – SIW cavity width, and C and ε_r are the velocity of light and the dielectric constant, respectively. The m, n, and p are mode indices.

A shorting pin is placed 3 mm from the left side and 8 mm from the top wall of the SIW cavity radiator to improve its bandwidth, as illustrated in Fig. 1 (a). The shorting pins establish a short circuit within the current path, leading to resonance. The increase in bandwidth is achieved by exciting two hybrid modes: resonance from the SIW cavity and shorting pins placed in close proximity to one another. The proposed SIW cavity-backed antenna features two slots designed to produce dual-band resonant frequencies: the outer ring Circular Split Ring Resonator(CSRR) slot and the inner ring CSRR slot. The integration of the inner and outer CSRR slots results in a reduction in series inductance while concurrently introducing series capacitance within the cavity circuit.

Antenna-2 has an outer ring CSRR slot positioned near the center of the side cavity, as illustrated in Fig. 1 (b). The resonant frequency of the CSRR outer ring slot can be calculated using the following equation:

$$f_r = \frac{c}{2L_1\sqrt{\varepsilon_r}} = \frac{3\times10^8}{2\times9\times10^{-3}\sqrt{4.4}} = 5.1\ GHz. \quad (2)$$

The placement of the CSRR outer ring slot significantly influences the current path of the first-order resonance mode (TE110), leading to a modified first-order mode (TE110) that shifts the resonant frequency from 3.4 GHz to approximately 2.5 GHz due to reactive loading, as illustrated in Fig. 1 (f). Consequently, the outer ring CSRR slot effectively regulates the lower resonance frequency (2.4 GHz) of the proposed antenna. The inner CSRR slot is removed from the center of the cavity within the outer ring CSRR in Antenna-3, as illustrated in Fig. 1 (c). The resonant frequency of the CSRR outer ring slot can be calculated using the following equation:

$$f_r = \frac{c}{2L_2\sqrt{\varepsilon_r}} = \frac{3\times10^8}{2\times18\times10^{-3}\sqrt{4.4}} = 2.48\ GHz. \quad (3)$$

The placement of the inner CSRR ring slot significantly influences the current path of the first order-resonance mode (TE110), resulting in a slight modification of the TE110 mode resonant frequency from

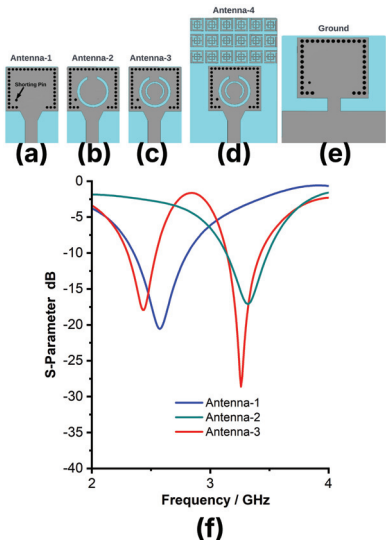


Fig. 1. (a)-(e) Development of the proposed SIW cavity-backed radiator, (f) Reflection coefficient of the proposed antenna.

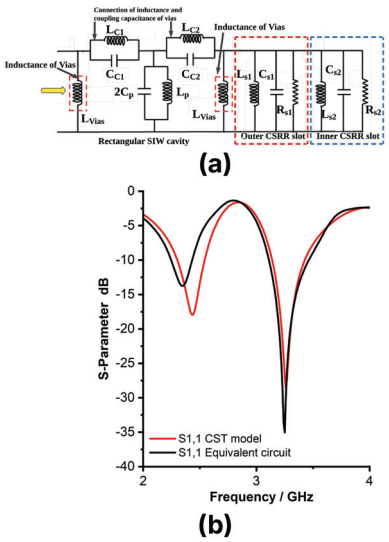


Fig. 2. Lumped equivalent circuit of proposed SIW cavity antenna, (b) Reflection coefficient of the proposed antenna for CST model and equivalent circuit model.

2.5 GHz to 2.41 GHz for WLAN applications. Additionally, the inner ring CSRR slot produces a second-order resonance mode (TE120) at 3.3 GHz as a result of the reactive loading from the CSRR slot, as illustrated in Fig. 1 (f). Ultimately, Antenna-3 exhibits a hybrid mode (TE110 and TE120) at frequencies of 2.4 and 3.3 GHz, suitable for WLAN and sub-6GHz 5G applications. The SIW cavity dual slot antenna is excited by a 50Ω feedline to function at frequencies of 2.4 GHz and 3.3 GHz. The lumped element equivalent circuit of the proposed SIW cavity-backed antenna is illustrated in Fig. 2 (a). The equivalent circuit of the SIW cavity, outer CSRR slot, and inner CSRR slot with conductor is illustrated in Fig. 2 (a). The equivalent circuit diagram illustrates that the metallic vias are influenced by inductance. The connection between the transmission line and the SIW cavity provides the coupling capacitance and coupling inductance. Additionally, the parasitic capacitance and inductance of the transmission line are represented by Cr and Lr. The outer CSRR ring slot is governed by the parallel lumped elements R, L, and C, while the inner CSRR ring slot is similarly governed by the parallel lumped elements R, L, and C.

The proposed SIW cavity-backed slot antenna is analyzed through the simulation of its lumped equivalent circuit using NI AWR Microwave Office software. The S-parameter of the lumped equivalent circuit of the proposed SIW cavity antenna is compared with the S-parameter of the CST model of the proposed SIW cavity antenna in Fig. 2 (b). Figure 2 (b) indicates that the proposed antenna operates at frequencies of 2.4 and 3.3 GHz for both the CST model and the equivalent circuit model. Thus, the performance of the CST model and the lumped equivalent circuit model is equivalent.

B. Metamaterial for gain enhancement and circular polarization conversion

The proposed antenna is intended for industrial WLAN and sub-6GHz 5G applications. Therefore, an antenna with standard gain and polarization is insufficient for this type of factory environment due to the high path loss. Therefore, an antenna with high gain and circular polarization is necessary for environments with significant path loss. This work introduces the metamaterial loading method to achieve high gain and circular polarization. The proposed Square Split Ring Resonator(SSRR) metamaterial unit cells are illustrated in Fig. 3 (a). The proposed SSRR metamaterial unit cell comprises an outer square ring stub and an inner square ring stub, which together provide an inductive effect, as depicted in Fig. 3 (a). Additionally, a 0.5 mm slot is removed at each corner of the outer square ring stub to generate a capacitive effect. As depicted in Fig. 3 (a), a 0.5-thickness stub is connected at the midpoint of both the inner and outer square ring stubs in vertical and horizontal orientations, which produces an inductive effect. The rectangular stub intersects both the inner and outer ring stubs at the center, as illustrated in Fig. 3 (a). In order to conduct the performance study of the metamaterial, the unit cell of the metamaterial is positioned between two waveguide ports. Additionally, multiple boundary conditions, including perfect electric conductor (PEC) and perfect magnetic conductor (PMC), were applied along the *xz axis*. The complete

characterization is conducted using CST Microwave Studio software. The incident electromagnetic waves typically propagate along the y-axis. The placement of the antenna regulates the y-axis stimulation of the metamaterial structure, subsequently influencing the transmission of electromagnetic waves via metamaterial.

The permeability, permittivity, and refractive index are essential parameters for accurately assessing the performance characteristics of the suggested SSRR metamaterial unit cell. The permittivity, permeability, and refractive index of the proposed metamaterial unit cell are illustrated in Figs. 4 (a-c). Figures 4 (a-c) demonstrate that the proposed metamaterial shows a negative real near-zero refractive index (NZRI) and epsilon-negative (ENG) characteristics at frequency ranges of 2.4 GHz and 3.3 GHz. The results clearly demonstrate that the proposed SSRR metamaterial shows near-zero properties within the operational bands (2.4 and 3.3 GHz). As a result, the designated operating frequency range (2.4 and 3.3 GHz) can be employed to enhance gain and adjust the polarization of the proposed antenna. The 6 × 3 array of modified Square Split Ring Resonator (SSRR) is depicted in Fig. 3 (b). The proposed design incorporates a metamaterial comprising a 6 × 3 array of modified Square Split Ring Resonator (SSRR) unit cells. It is printed in front of the antenna in the y-direction as depicted in Fig. 1 (d). The periodic distance of 0.5 mm is maintained in both the x and y directions between two SSRR unit cells.The y-axis is the orientation of an electromagnetic wave that is propagating normally. The y-axis corresponds to the normal path of the electromagnetic wave's incidence. Electromagnetic interactions within the modified SSRR metamaterial unit cell, triggered by the incident waves, produce resonance in the transmitted and reflected waves. Since the electromagnetic wave enters the modified SSRR metamaterial unit cells via the SIW cavity-supported antenna and travels in that direction, the y-axis makes it easier to excite the modified SSRR metamaterial unit cell.

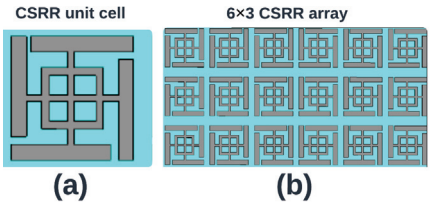


Fig. 3. (a-b) Proposed SSRR metamaterial unitcell.

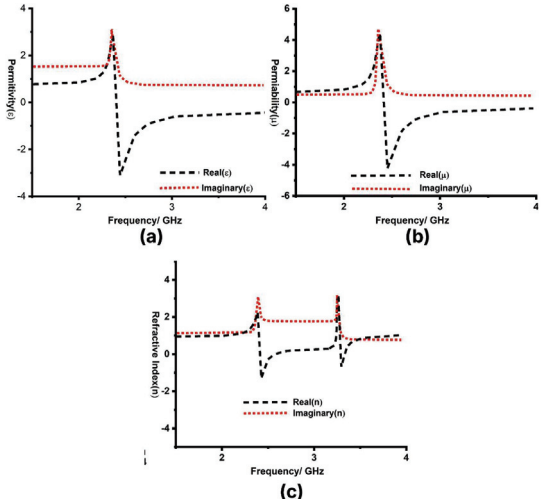


Fig. 4. (a) Permittivity, (b) Permeability, and (c) Refractive index of SSRR unitcell.

The antenna's gain, with and without metamaterial, is illustrated in Fig. 5 (a). Figure 5 (a) reveals that the suggested antenna demonstrates gains of 3.2 dB and 2.6 dB at frequencies of 2.4 GHz and 3.3 GHz, respectively. The suggested antenna demonstrates a gain of 6.1 dB at 2.4 GHz and 6.5 dB at 3.3 GHz, following the integration of the metamaterial unit cells. The gain of the proposed antenna with SSRR metamaterial loading has been significantly improved without modifying the resonant frequencies, as illustrated in Fig. 5 (a). Moreover, the metamaterial generally has the capacity to modify polarization. The suggested metamaterial unit cells are situated in the y-direction in front of the antenna, facilitating the conversion of linear polarization into circular polarization. The suggested metamaterial functions as a linear-to-circular polarization converter at the 2.4 and 3.3 GHz bands, as evidenced by the axial ratio characteristic illustrated in Fig. 5 (b). The axial ratio of the proposed antenna exceeds three decibels at all operational frequencies (2.4 and 3.3 GHz), as illustrated in Fig. 5 (b).

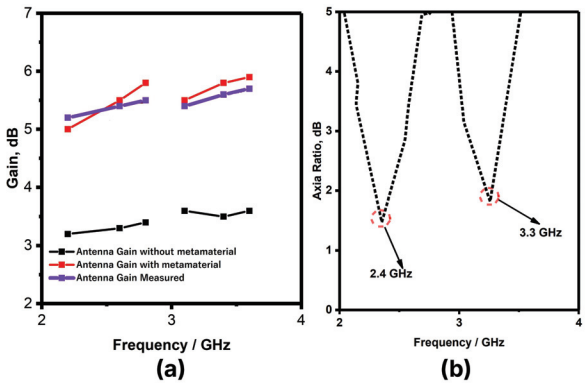


Fig. 5. (a) Gain of the the proposed radiator with and without metameterial, and (b) Axial ratio of the proposed radiator with and without metamaterial.

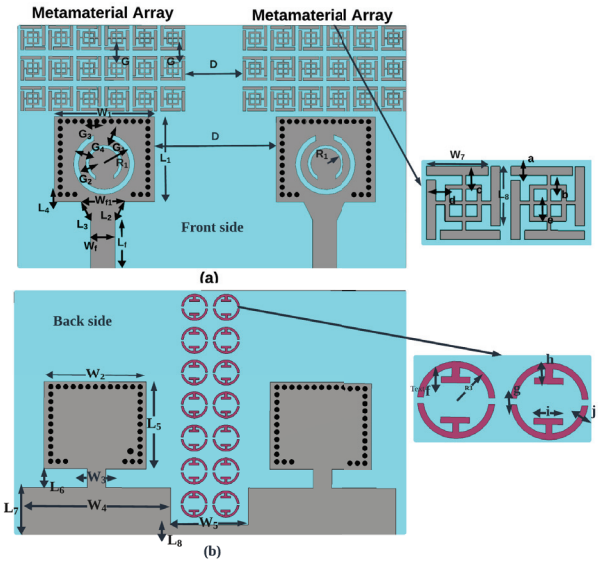


Fig. 6. Proposed MIMO antenna without and with SIW cavity backed parasitic stub decoupling structure (a) Front side, and (b) Back side of the substrate.

C. Two-port SIW cavity-backed high-gain circularly polarized MIMO antenna design

For the purpose of facilitating communication in an environment that has a considerable path loss due to scattering, the antenna ought to have a high channel capacity, a high gain, and a good link reliability. This led us to propose the two-port SIW cavity-backed gain improved circularly polarized metamaterial-loaded dual-band MIMO antenna as a potential solution for usage in WLAN and sub-6GHz 5G applications in industrial environments. As depicted in Fig. 6, the SIW cavity-backed gain increased the circularly polarized metamaterial-loaded dual-band MIMO antenna that has been presented is designed by positioning two SIW-based metamaterial-loaded radiators in close proximity to each other so that they share a common ground plane. The adjacent positioning of the proposed radiator enhances the mutual connection between the suggested SIW cavity-backed radiators. Nevertheless, the mutual coupling among the radiators requires further enhancement to optimize MIMO performance. There is a necessity for a distinct decoupling structure between the proposed SIW cavity-backed radiators. An EBG structure can inhibit the transmission of electromagnetic waves within specific frequency ranges, termed the band gap. The EBG exhibits band stop filter characteristics across the band gap, enabling its application as a decoupling structure in MIMO antennas. Therefore, to diminish the reciprocal coupling between the radiators in the proposed MIMO antenna, a unique EBG decoupling structure is presented in Fig. 7 (a). The unit cell of the proposed modified CSRR EBG structure is seen in Fig. 7 (a). The proposed EBG decoupling structure is constructed by using modified CSRR ring stubs as depicted in Fig. 7 (a). As depicted Fig. 7 (a), two T-shaped stubs are connected along with modified CSRR of the EBG unit cell.

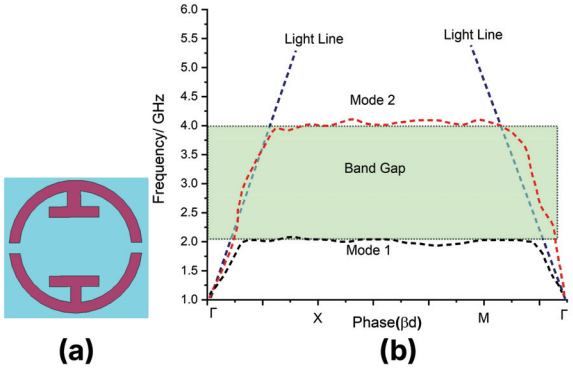


Fig. 7. (a) Proposed modified Circular Split Ring Resonator EBG unit cell, and (b) Dispersion diagram of the proposed CSRR EBG unit cell.

The characterization of the proposed EBG unit cell is conducted using CST Studio software. To effectively analyze the features of the proposed EBG unit cell, the dispersion diagram of the EBG can be utilized. The dispersion diagram of the proposed EBG unit cell is illustrated in Fig. 7 (b). Figure 7 (b) indicates that the proposed EBG unit cell functions as a stop band within the frequency range of 2.2 to 4 GHz. Consequently, throughout the 2.2-4 GHz frequency range, the EBG unit cell functions as a stop-band filter. The electromagnetic waves propagate from one radiator to another while both radiators are simultaneously energized. To alleviate this issue, the suggested 7x2 EBG unit cells are positioned on the rear side of the substrate, above the ground plane,

Table 1: Dimension of the proposed antenna

P=D	P=D	P=D	P=D
L=40	LG2=1.9	W2=1	WG1=3
L1=12.1	LG3=12.1	W3=2.3	WG2=1.5
L2=5.5	LG4=1	W4=7.3	WG3=5
L3=1.8	LG5=1.2	W5=7.5	WG4=3.5
L5=5	LG6=18	W6=3	WG7=12
L6=8.3	LG7=y2	W7=1	WG8=y20
L7=1.5	LG8=2	W8=1.5	WG10=y1
L8=1.4	LG9=1	W9=1.5	G=1
LG=24.5	W=35	W10=1.5	R=0.4
LG1=1.4	W1=y10	WG=44	D1=2
D3=1.2	D4=0.5	D5=y1.2	D6=1
D7=y0.5	D8=0.5	D9=0.5	D10=0.5
D11=3	E1=2.1	E2=2.1	E3=0.6
E4=0.6	E5=0.5	E6=0.5	E7=y0.3

situated between the two SIW cavity-backed radiators, as illustrated in Fig. 6 (b). The positioning of the EBG unit cell facilitates the absorption of surface current by the EBG unit cells when both radiators are simultaneously activated. This leads to a substantial decrease in mutual coupling among the radiators in the proposed MIMO antenna at the 2.4 and 3.3 GHz frequencies. The dimension of the proposed MIMO antenna is depicted in Table 1.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Reflection coefficient and mutual coupling

The reflection coefficient and mutual coupling of the proposed MIMO antenna are measured using a VNA, as illustrated in Fig. 8. The simulated and measured reflection coefficients of the proposed MIMO antenna are compared in Fig. 9 (a).

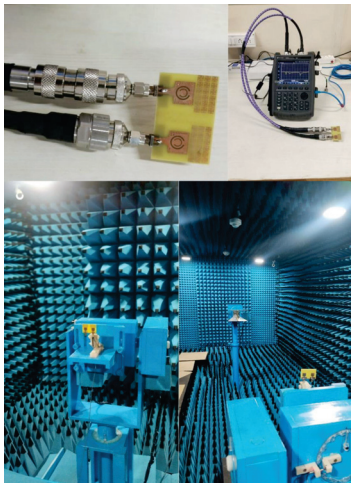


Fig. 8. Photograph of the fabricated antenna and measurement setup of the proposed MIMO antenna.

Figure 9 (a) indicates that the suggested MIMO antenna has a reflection coefficient of less than -10 dB at 2.4 and 3.3 GHz. The mutual coupling of the proposed MIMO antenna, both simulated and measured, is compared in Fig. 9 (b). Figure 9 (b) indicates that the suggested MIMO antenna exhibits mutual coupling of less than -15 dB at 2.4 and 3.3 GHz. A high degree of agreement can be found between the results of the simulated and measured reflection coefficients as well as the mutual coupling studies.

B. Radiation pattern and efficiency

The radiation pattern of the proposed MIMO is assessed in an anechoic room, as depicted in Fig. 8. The cross and co-polarization in E- and H-planes at 2.4 and 3.3 GHz is seen in Figs. 10 (a-d). Figures 10 (a-d), illustrates that the proposed antenna exhibits

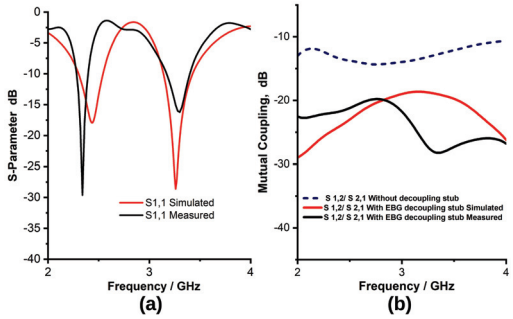


Fig. 9. (a)-(b) Simulated and measured reflection coefficient and mutual coupling of the proposed MIMO antenna.

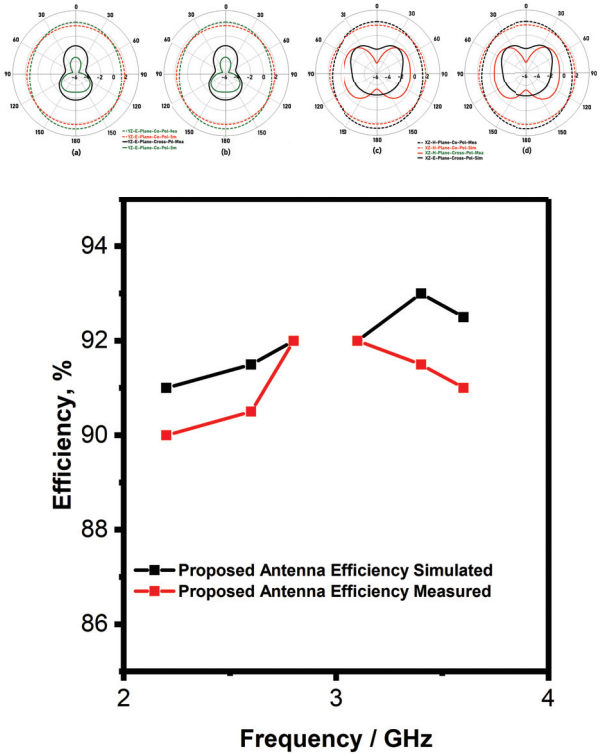


Fig. 10. (a)-(d) Radiation pattern of the proposed antenna at 2.4, and 3.3 GHz in E- and H-plane, and (e) Efficiency of the proposed antenna at 2.4 and 3.3 *GHz*.

an omnidirectional radiation pattern at 2.4 and 3.3 GHz. Due to the metamaterial loading and the SIW cavity, cross-polarization and backlobe radiation are suppressed, resulting in the proposed antenna exhibiting an excellent co-polarization radiation pattern at 2.4 and 3.3 GHz. There is a high degree of concordance between the results of the simulated and measured radiation patterns.

A representation of the simulated and observed efficiency of the proposed MIMO antenna can be found in Fig. 10 (e). According to the data presented in Fig. 10 (e),

the suggested MIMO demonstrates an efficiency that surpasses 90% at all operational bands (2.4 and 3.3 GHz).

IV. DIVERSITY PERFORMANCE OF PROPOSED MIMO ANTENNA

A. Diversity gain

Diversity gain, which quantifies the improvement in system performance due to the application of diversity techniques, is a critical metric calculated using equation (4). The performance of an enhanced system is indicated by an increased diversity gain value. The optimal number for MIMO antenna performance should be approximately 10 to achieve expected results.

$$DG = 10\sqrt{1-|ECC|^2}. \tag{4}$$

The diversity gain of the proposed MIMO antenna is illustrated in Fig. 11 (a). From Fig. 11 (a), it is evident that the proposed MIMO antenna demonstrates a diversity gain value of approximately 10 at frequencies of 2.4 and and 3.3 GHz.

B. Envelope Correlation Coefficient (ECC)

The ECC, is a statistical measure that is used to quantify the degree of correlation that occurs between the components of the MIMO antenna:

$$ECC = \frac{\left|\iint \overrightarrow{F_1}(\theta,\phi)\overrightarrow{F_2}(\theta,\phi)d\Omega\right|^2}{\iint\left|\overrightarrow{F_1}(\theta,\phi)\right|d\Omega^2 \iint\left|\overrightarrow{F_2}(\theta,\phi)\right|d\Omega^2} \tag{5}$$

$$ECC = \frac{|S_{11}{}^*S_{12}+S_{21}{}^*S_{22}|^2}{\left(1-\left(|S_{11}|^2+|S_{21}|^2\right)\right)\left(1-\left(|S_{22}|^2+|S_{12}|^2\right)\right)}. \tag{6}$$

The determination can be achieved through scattering parameters by applying equation (5) in a lossless environment, characterized by a uniform distribution of power across antenna elements. This equation is applicable solely in scenarios where there is no loss present. Consequently, the ECC can be estimated based on the emitted far-field by utilizing equation (6). For uncorrelated MIMO antennas, the optimal value of the ECC is zero. However, for practical MIMO antennas, the ECC value must not exceed 0.5. Figure 11 (b) illustrates the proposed MIMO antenna based on the ECC design. The data presented in Fig. 11 (b) indicates that the proposed MIMO antenna demonstrates an ECC value of less than 0.04 at frequencies of 2.4 and 3.3 GHz.

C. TARC

An accurate evaluation of the correlation between the components of the MIMO antenna can be accomplished with the use of a statistical measure known as TARC. It is possible to compute the TARC of the MIMO antenna system for two ports by utilizing equation (7). In order to get optimal performance, the TARC of the

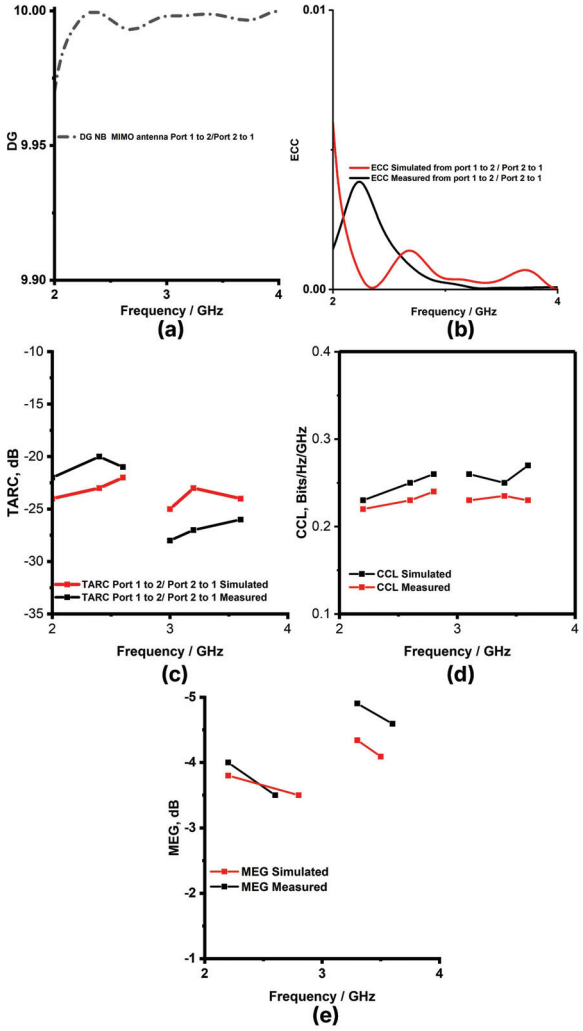


Fig. 11. (a) ECC, (b) DG, (c) TARC, (d) CCL, and (e) MEG of the proposed MIMO antenna.

MIMO antenna must be lower than 0 dB:

$$TARC = \sqrt{\frac{|(S_{11}+S_{12}e^{j\theta})|^2+|(S_{21}+S_{22}e^{j\theta})|^2}{2}} \tag{7}$$

Figure 11 (c) illustrates the TARC of the MIMO antenna. The data presented in Fig. 11 (c) indicates that the proposed MIMO antenna achieves a TARC of less than -15 dB. As a result, the proposed MIMO antennas demonstrate a reduced correlation between the radiating elements.

D. CCL

The primary function of MIMO antennas is to enhance the capacity of the communication channel. Nevertheless, losses are a consequence of the correlation that exists between the elements of the antenna. For the purpose of estimating the highest practical limit for signal transmission while simultaneously reducing sig-

Table 2: Comparison of the proposed work with existing work

P = D	P = D	P = D	P = D
L = 40	LG2=1.9	W2=1	WG1=3
L1=12.1	LG3 = 12.1	W3=2.3	WG2 = 1.5
L2 = 5.5	LG4=1	W4 = 7.3	WG3 = 5
L3 = 1.8	LG5 = 1.2	W5 = 7.5	WG4 = 3.5
L5 = 5	LG6=18	W6 = 3	WG7 = 12
L6 = 8.3	LG7 = 2	W7 = 1	WG8 = 20
L7 = 1.5	LG8=2	W8 = 1.5	WG10=1
L8 = 1.4	LG9 = 1	W9 = 1.5	G = 1
LG = 24.5	W = 35	W10 = 1.5	R = 0.4
LG1=1.4	W1=10	WG=44	D1=2
D3 = 1.2	D4 = 0.5	D5 = 1.2	D6=1
D7 = 0.5	D8 = 0.5	D9 = 0.5	D10 = 0.5
D11=3	E1 = 2.1	E2 = 2.1	E3 = 0.6
E4 = 0.6	E5 = 0.5	E6 = 0.5	E7 = 0.3

nificant loss, a metric known as channel capacity loss (CCL) is utilized. The calculation of CCL can be carried out by utilizing equation (8). In real applications, the CCL should be kept at a level that is lower than 0.4 Bits/Hz/Sec:

$$CCL = -\log_2 \det\left(\psi^R\right), \quad \left(\psi^R\right) = \begin{bmatrix} \rho_{ii} & \rho_{ij} \\ \rho_{ji} & \rho_{jj} \end{bmatrix}. \tag{8}$$

Figure 11 (d) illustrates a representation of the CCL of the MIMO antenna. The data presented in Fig. 11 (d) indicates that the proposed MIMO antenna achieves a CLL of less than 0.4 Bit/Hz/Sec at frequencies of 2.4 and 3.3 GHz.

E. MEG

The MEG statistic is another important statistic that is used to evaluate the performance of MIMO antennas. This statistic calculates the average signal strength that is received by each radiator within the antenna. The mathematical expression known as MEGi represents the ratio of the mean power that is received by the ith element to the mean power that is incident on the jth element from the same source. It is possible to calculate MEG by using equation (9), and the MEG value ought to be lower than 3 dB in order to achieve the best possible MIMO performance:

$$\mathrm{MEG_i} = 0.5\left[1 - \sum_{\mathrm{i=1}}^{\mathrm{N}} \left|\mathrm{S_{ij}}\right|^2\right]. \tag{9}$$

The MEG of the MIMO antenna is depicted in Fig. 11 (e). The data presented in Fig. 11 (e) indicates that the proposed MIMO antenna exhibits a MEG of less than -3 dB at frequencies of 2.4 GHz and 3.3 GHz.

V. COMPARISON OF THE PROPOSED WORK WITH EXISTING WORK

Table 2 illustrates a comparison of the performance metrics between the proposed MIMO antenna and the existing MIMO antennas. The data presented in Table 2 indicates that the recommended MIMO antenna outperforms the currently utilized antennas regarding size, innovative methodology, and gain.

VI. CONCLUSION

In this work, a SIW cavity-backed gain-enhanced circularly polarized metamaterial-loaded dual-band MIMO antenna is designed for WLAN and sub-6 GHz 5G applications. The dual operating bands are enabled by placing a modified CSRR-shaped slot at the center of the SIW cavity-backed rectangular radiator. Additionally, the antenna incorporates 6×3 modified SSRR unit cells strategically located in front of the intended radiators along the *y*-axis. This arrangement enabled circular polarization and enhanced the gain of the proposed radiator at 2.4 GHz and 5 GHz. The metamaterial loading of the proposed antenna yields a gain of 5.5 dB at 2.4 GHz and 5.4 dB at 5 GHz. The cavity-backed configuration of the suggested antenna yields a unidirectional radiation pattern, contributing to its efficiency surpassing 90% at frequencies of 2.4 and 5 GHz. Further, the implementation of a CSRR EBG decoupling structure reduces the mutual coupling between the radiators. The antenna exhibits exceptional diversity performance, as evidenced by various performance metrics: an ECC of 0.001, a DG of 9.98, a CCL of < 0.4 bits $/\mathrm{S}/\mathrm{Hz}$, a MEG of < 3 dB, and a TARC of < -20 dB. The experimental validation of the system confirms its intended functionality, showing a strong correlation between the simulated and measured results.

REFERENCES

[1] M. Kaushik, J. K. Dhanoa, and M. K. Khandelwal, "Partially omnidirectional and circularly polarized MIMO antenna covering sub-6-GHz band for 5G fast plan," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 13, no. 9, pp. 1443-1450, 2023.

[2] A. Kumar, C. Rai, and P. Chaurasia, "Realization of miniaturized triple-band four-port stacked MIMO antenna for WLAN applications at 2.9/5.0/5.9 GHz bands," *AEU - International Journal of Electronics and Communications*, vol. 150, p. 154216, 2022.

[3] R. Khandelwal, "Metamaterial based circularly polarized four-port MIMO diversity antenna embedded with slow-wave structure for miniaturization and suppression of mutual coupling," *AEU - International Journal of Electronics and Communications*, vol. 121, p. 153241, 2020.

[4] U. Ullah, M. Al-Hasan, S. Koziel, and I. B. Mabrouk, "Series-slot-fed circularly polarized multiple-input–multiple-output antenna array enabling circular polarization diversity for 5G 28 GHz indoor applications," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 9, pp. 5607-5616, 2021.

[5] S. Virothu and M. Satya Anuradha, "Flexible CP diversity antenna for 5G cellular Vehicle-to-Everything applications," *AEU - International Journal of Electronics and Communications*, vol. 152, p. 154248, 2022.

[6] M. Tiwari, K. Afroz, and M. Fatima, "Wideband metasurface loaded circularly polarized MIMO microstrip antenna with high isolation," *Microwave Review*, vol. 28, pp. 22-27, 2022.

[7] J. Wang, Y. Cheng, H. Luo, F. Chen, and L. Wu, "High-gain bidirectional radiative circularly polarized antenna based on focusing metasurface," *AEU - International Journal of Electronics and Communications*, vol. 151, p. 154222, 2022.

[8] S. Luo, Y. Li, Y. Xia, and L. Zhang, "A low mutual coupling antenna array with gain enhancement using metamaterial loading and neutralization line structure," *Appl. Comput. Electromagn. Soc. J*, vol. 34, pp. 411-418, 2019.

[9] M. Y. Li, Z. Q. Xu, Y. L. Ban, C. Y. D. Sim, and Z. F. Yu, "Eight-port orthogonally dual-polarised MIMO antennas using loop structures for 5G smartphone," *IET Microw. Antennas Propag*, vol. 11, pp. 1810-1816, 2017.

[10] S. Mukherjee and A. Biswas, "Design of planar high-gain antenna using SIW cavity hybrid mode," *IEEE Trans. Antennas Propag*, vol. 66, no. 2, pp. 972-977, 2018.

[11] A. Kumar, Divya Chaturvedi, and S. Raghavan, "SIW cavity-backed circularly polarized square ring slot antenna with wide axial-ratio bandwidth," *AEU - International Journal of Electronics and Communications*, vol. 94, pp. 122-127, 2018.

[12] G. Q. Luo, Z. F. Hu, W. J. Li, X. H. Zhang, L. L. Sun, and J. F. Zheng, "Bandwidth-enhanced low-profile cavity-backed slot antenna by using hybrid SIW cavity modes," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 4, pp. 1698-1708, 2012.

[13] S. Yun, D. Kim, and S. Nam, "Bandwidth enhancement of cavity backed slot antenna using a via-hole above the slot," *IEEE Antennas Wireless Propagation Letter*, vol. 11, pp. 1092-1095, 2012.

[14] H. Liu, H. Tian, L. Liu, and L. Feng, "Co-design of wideband filtering dielectric resonator antenna with high gain," *IEEE Trans. CircuitsSyst. II Express Briefs*, vol. 69, pp. 1064-1068, 2022.

[15] C. Shi, J. Zou, J. Gao, and C. Liu, "Gain enhancement of a dual-band antenna with the FSS," *Electronics*, pp.28-34, 2022.

[16] B. A. F. Esmail, S. Koziel, and D. Isleifson, "Metamaterial-based series-fed antenna with a high gain and wideband performance for millimeter-wave spectrum applications," *Electronics*, vol. 12, no. 23, pp. 4836-4841 2023.

[17] B. A. F. Esmail, S. Koziel, and A. Pietrenko-Dabrowska, "Wideband high-gain low-profile series-fed antenna integrated with optimized metamaterials for 5G millimeter wave applications," *Scientific Report*, vol. 14, pp. 185-192, 2024.

[18] A. Gorai, A. Deb, and J. R. Panda, "Millimeter wave/5G multiband SIW antenna with metasurface loading for circular polarization and bandwidth enhancement," *Journal of Infrared Milli Terahertz Waves*, vol. 43, pp. 366-383, 2022.

[19] F. Ez-Zaki, "Double negative DNG metamaterial-based Koch fractal MIMO antenna design for sub-6 GHZ V2X communication," *IEEE Access*, vol. 11, pp. 77620-77635, 2023.

[20] I. E. Lamri and A. Mansoul, "Design of CPW-fed dual-band four-element MIMO microstrip patch antenna for WLAN/WiMAX applications," *Institute of Electrical and Electronics Engineers Inc*, pp. 83-86, 2023.

[21] V. Kikan, A. Dagar, S. Singh, S. Singh, N. C. Deo, A. Kumar, and M. Sharma, "A four-port novel inset-fed, rectangular MIMO-antenna designed for 2.40 GHz Bluetooth & Wi-Fi applications," in 2023 *Second International Conference on Electrical, Electronics, Information Communication Technology (ICEEICT)*, pp. 1–6, 2023.

[22] A. Pal and V. S. Tripathi, "Quad-element MIMO antenna with diverse radiation pattern characteristics and enhanced gain for 5.9 GHZ V2X communications," *AEU-Int. J. Electron. Commun*, vol. 176, p. 155119, 2024.

[23] X. Mao, Z. Zhang, and J. Wang, "Dual-polarized reconfigurable MIMO decoupled antenna design using characteristic mode analysis," *IEEE conference*, pp. 35–41, 2023.

[24] A. Kumar, P. V. Neeraj, and S. C. Padhy, "Four port MIMO antenna for IoT applications in public safety band and sub-6 GHz TDD 5G band," *Eng. Res. Express*, vol. 6, p. 015309, 2024.

[25] R. Khan, W. T. Sethi, and W. A. Malik, "Enhancing gain and isolation of a quad-element MIMO antenna array design for 5G sub-6 GHz applications assisted with characteristic mode analysis," *Scientific Report*, vol. 14, pp. 108–117, 2024.

Infant Leo S received the Bachelor of Engineering degree in Electronics and Communication Engineering and Master of Engineering degree in Communication Systems from the Anna University, Chennai, Tamil Nadu, India, in 2015 and 2019, respectively, and is currently pursuing the Doctor of philosophy from the National Institute of Technology Tiruchirappalli, Tamil Nadu, India. He is currently working as an Assistant Professor in the Department of Electrical, Electronics and Communication Engineering at Galgotias University in Greater Noida, Gautam Buddh Nagar, Uttar Pradesh. He has been interested in RF Microwave antennas circuit design and electromagnetics for more than 8 years.

G. Aloy Anuja Mary received the B.E. and M.E. degrees in Electronics Engineering from the Manonmaniam and Anna University in 2003 and 2005, respectively, and Ph.D from Anna University, Chennai, Tamil Nadu, India. She is currently Professor in ECE at Vel tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology,Chennai. She has been interested in the WirelessCommunication and Network Security for more than 15 years.

A. Syed Mazhar received the B.E, M.E and Ph.D. (Pursuing) in Electronics and Communication Engineering from the Anna University, Chennai, Tamil Nadu India, in 2008 and 2012, respectively, and an MBA in Project Management from the Alagappa University, Karaikudi, Tamil Nadu, India. He is currently working as an Assistant Professor Gr-3 in ECE at Galgotias University, Greater Noida, U.P, India. He has been interested in medical electronics, microprocessor and microcontroller and satellite communications for more than 13 years.

Satyam Mishra received the Bachelor of Technology degree in Electronics and Communication Engineering from the Dr.A.P.J.Abdul Kalam Technical University at Uttar Pradesh, India, in 2017 and Master of Technology degree in Signal Processing and Digital design from the Delhi Technological University at Delhi, India, in 2022 and is currently pursuing the Doctor of philosophy from the National Institute of Technology Delhi, India. He is currently working as an Assistant Professor in the Department of Electrical, Electronics and Communication Engineering at Galgotias University in Greater Noida, Gautam Buddh Nagar, Uttar Pradesh. He has been interested in the numerical aspects of antenna design and electromagnetic Machine Learning for more than 5 years.

G. Jothi received the B.E. and M.E. degrees in Computer Science Engineering from Anna University, Chennai, Tamil Nadu India. She is an adjunct professor in ECE at Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai.