IEEE Open Journal of
Antennas and Propagation

Received 12 January 2021; revised 17 February 2021; accepted 18 February 2021. Date of publication 24 February 2021; date of current version 5 March 2021.

Digital Object Identifier 10.1109/OJAP.2021.3061936

Low-Rank Matrix Factorization Method for Multiscale Simulations: A Review

MENGMENG LI [1] (Senior Member, IEEE), DAZHI DING [1], ALEXANDER HELDRING [2], JUN HU [3] (Senior Member, IEEE), RUSHAN CHEN [1] (Senior Member, IEEE), AND GIUSEPPE VECCHI [4] (Fellow, IEEE)

[1]Department of Communication Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
[2]Antenna Laboratory, Department of Signal Processing and Telecommunications, Universitat Politecnica de Catalunya, 08034 Barcelona, Spain
[3]School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[4]Antenna and EMC Laboratory (LACE), Politecnico di Torino, 10125 Turin, Italy

CORRESPONDING AUTHOR: D Ding (e-mail: dzding@njust.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 61871222, Grant 62025109, Grant 61890541, Grant 61931021, and Grant 61721001; and in part by the Spanish Government under Project TEC 2017-83343-C4-2-R and Project MDM-2016-0600.
This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/OJAP.2021.3061936.

ABSTRACT In this paper, a review of the low-rank factorization method is presented, with emphasis on their application to multiscale problems. Low-rank matrix factorization methods exploit the rank-deficient nature of coupling impedance matrix blocks between two separated groups. They are widely used, because they are purely algebraic and kernel free. To improve the computation precision and efficiency of low-rank based methods, the improved sampling technologies of adaptive cross approximation (ACA), post compression methods, and the nested low-rank factorizations are introduced. O(N) and O ($N\log N$) computation complexity of the nested equivalence source approximation can be achieved in low and high frequency regime, which is parallel to the multilevel fast multipole algorithm, N is the number of unknowns. Efficient direct solution and high efficiency preconditioning techniques can be achieved with the low-rank factorization matrices. The trade-off between computation efficiency and time are discussed with respect to the number of levels for low-rank factorizations.

INDEX TERMS Integral equation, method of moments, low-rank, multiscale.

I. INTRODUCTION

INTEGRAL equation methods are the preferred methods when modeling and simulating large and multiscale problems due to their high computation precision and small number of unknowns. However, the method of moments (MoM) discretized integral equations will lead to dense impedance matrices [1], [2]. The computation time and memory requirements have complexity $O(N^3)$ and $O(N^2)$ respectively, where N is the number of unknowns, when a straightforward direct solver is used.

To enhance the computational performance of MoM, there are mainly three kinds of fast solvers based on different computation acceleration techniques:

(1) Sparsification of the impedance matrix by using specialized basis functions such as wavelet expansions [3] and impedance matrix location [4]. Both computation time and memory requirements can be saved due to the reduced number of nonzero elements of the impedance matrix.

(2) Reduction of the matrix size through compression of the impedance matrix by using macro basis functions [5]. The characteristic basis functions [6], [7], synthetic basis functions [8], [9], and sub-entire-domain (SED) basis functions [10], [11] are proposed for the computation of large-scale problems. Recently, similar to the macro basis functions, dominant characteristic modes [12] are extracted to reduce the dimensions of MoM impedance matrix for the analysis of antenna arrays [13].

(3) To evaluate the dense off-diagonal sub-blocks of the impedance matrix with multilevel fast multipole algorithm

(MLFMA) [14], [15], FFT-based method [16]–[19], or low-rank matrix factorization methods [20]–[32].

The MLFMA reduces the computational complexity to $O(N \log N)$ when the targets are discretized with a fixed mesh density with respect to the wavelength. However, the MLFMA implementation depends on a priori knowledge of the integral kernel (Green's function). For the FFT based methods, the current basis functions defined on discretized cells are projected on to regular grids by interpolation and projection, to reproduce the same fields as the original currents [16]–[19]. The FFT based methods show better performance for the volume integral equation than for the surface integral equation, as in the former case the unknowns are distributed throughout the entire space of the regular grids. For surface integral equations, especially for the evaluation of inhomogeneous and multiscale problems, the regular grids in empty computation domain worsens the computation performance [41]. Moreover, the near-field interactions cannot be evaluated correctly with interpolation and projection, therefore pre-corrected computations are required. The multilevel recursive interpolation and subdomain FFT are introduced in multilevel FFT [42], [43], where $O(N)$ computational complexity can be achieved for static capacitance extraction [42].

The low-rank factorization methods exploit the rank-deficient nature of coupling impedance matrix blocks between two separated groups. The whole impedance matrix of MoM is full-rank, while the off-diagonal matrix blocks are low-rank due to the basis functions over sampling than Nyquist limit for "far" couplings [44]. The typical use of low-rank factorization techniques in the MoM consists of a block decomposition of the impedance matrix, followed by compression of those blocks that represent interactions between well-separated regions of the target geometry. This yields an approximate representation of the impedance matrix that can be used for fast matrix-vector products in iterative Krylov-subspace solvers such as GMRES or BiCGStab [45].

An important appeal of the low-rank factorization methods over other fast methods is that they are purely algebraic and kernel free; the low-rank approximation matrices are usually obtained by algebraic operations, such as QR decomposition, singular value decomposition (SVD), and adaptive cross approximation (ACA) [23], [24]. As a result, the low-rank factorization methods can be easily employed as black box fast solvers for accelerating existing MoM codes.

Among the best-known algorithms for fast low-rank factorization of off-diagonal impedance matrix blocks is the ACA algorithm. Originally proposed by Bebendorf in 2000 [23] for problems with non-oscillating kernels, it was first successfully applied to problems in computational electromagnetics (CEM) in 2005 by Zhao *et al.* [24]. The algorithm is purely algebraic; it can be seen as a truncated LU decomposition of a low-rank matrix sub block $\mathbf{Z}$. If $\mathbf{Z}$ has dimensions $m \times n$, ACA generates an approximation of $\mathbf{Z}$ as a product $\mathbf{UV}$, where $\mathbf{U}$ is $m \times r$ and $\mathbf{V}$ is $r \times n$, r being the rank that is necessary to approximate $\mathbf{Z}$ with a predefined precision τ. The computational cost is proportional to $r^2(m+n)$.

TABLE 1. Performance comparisons of the low-rank factorization methods, FFT based methods, and MLFMA.

Characteristics	Performance
Easy implementation	Low-rank methods > FFT methods > MLFMA
Kernel independence	Low-rank methods > FFT methods > MLFMA
Computation efficiency	MLFMA > FFT methods > Low-rank methods

Although the gain in efficiency with respect to direct inversion of the full impedance matrix is impressive, the degrees of freedom in the interaction between separate regions of the target and hence the rank of the corresponding off-diagonal impedance matrix block grows with the electrical size, initially approximately proportional to the frequency but for asymptotically large frequency it will grow with the frequency squared [49]. As a consequence, the computational burden of the ACA compression will eventually grow proportionally to $O(N^3)$ and the storage to $O(N^2)$. This severely limits the applicability of ACA to electrically large problems.

Table 1 shows the performance comparisons of the low-rank factorization methods with FFT based methods and MLFMA, the low-rank factorization methods are the easiest to implement but their computational performance is worse than FFT methods and MLFMA. Different from the MLFMA, the low-rank decomposition processes are time consuming, since recursive relations between neighboring levels are difficult to find for traditional low rank methods [20]–[32], due to their fully algebraic nature. As a result, despite their obvious advantages (kernel independence, ease of implementation, well controlled accuracy), ACA compressed iterative solvers eventually lose out to the famed $O(N \log N)$ complexity of competing algorithms such as MLFMA [14], [15]. Various approaches are currently under investigation by different research groups with the aim of bringing down the complexity of the ACA and other low rank methods.

The bottleneck of traditional low-rank methods can be identified in the fact that, even for a multilevel algorithm, the low rank approximations must be explicitly computed and stored at each level, which in turn worsens the setup time and storage requirements. A first step in alleviating the above limitation is to define the sub regions in a recursive-, hierarchical- or multiscale manner, and adopt a judiciously chosen *admissibility criterion* at every *scale* or *level* in the hierarchy. The admissibility criterion determines whether a matrix block corresponding to a given level is ACA compressed or not. If it is not, then its *children* (the sub blocks inside of it) are either ACA compressed or not, again depending on the admissibility criterion. Since the rank not only depends on the electrical size but also on the mutual distance between regions, the admissibility criterion can be optimized with

respect to these two parameters, yielding $O(N^{3/2})$ complexity both for computation time and memory requirements, as shown in [49]. Matrices constructed according to this procedure of hierarchical subdivision and block wise compression according to an admissibility criterion have been known by mathematicians since 1999 as H-matrices [22].

An alternative option exploits the concept, first proposed in [21], of a *butterfly* decomposition of the low-rank interaction matrix of well-separated sub-groups [46], [47]. As noted above, the rank r asymptotically depends linearly on the group size n, hence the growth of complexity with the electrical size of the problem. However, as demonstrated in [21], the interaction matrix can be decomposed into a sequence of length $O(\log n)$ of sparse matrices, each with no more than $O(n)$ non-zero elements. Replacing all the ordinary $O(rn)$ low-rank blocks in the H-matrices format with butterfly decompositions results in an $O(N\log^2 N)$ representation. The bottle-neck in this approach is the computational cost of constructing the butterfly decomposition. In [21] and [27], this was done with the help of "auxiliary" sources. Finding optimum sets of such auxiliary sources in terms of accuracy and efficiency is presently the main obstacle for this approach to be competitive. In [28], an entirely algebraic algorithm was proposed, the multilevel ACA (MLACA) that constructs the butterfly decomposition using only ACA compression. This algorithm does indeed achieve the expected high degree of compression and $O(N\log^2 N)$ memory scaling. However, unfortunately the complexity of the matrix compression is $O(N^2 \log N)$. Recently, a parallel hierarchical blocked ACA is proposed, where reduced computational complexity for low-rank factorizations can be achieved [48].

The MLACA can be considered a post compression technique. Other post compression technique are SVD [27], [50]–[52], multilevel ACA [28], [53], [54], multilevel matrix compression methods [30], and multilevel simply sparse method (MLSSM) [55], [56]. They are applied to the low-rank approximation matrices to achieve further compression. The post compression techniques reduce the memory requirements and matrix-vector product time significantly due to a relative smaller rank with respect to standard low-rank factorization methods. However, the matrix factorization setup time to construct the low-rank approximation is typically increased [28]. Therefore they are the preferred choice when the solution step dominates, for example for computations with multiple right-hand vectors.

To enhance the computational performance of the low-rank factorization methods, two kinds of low-rank factorization methods for Green's functions and impedance matrices have been developed recently for realistic multiscale simulations. For the H^2 method, Lagrange interpolation methods are employed to approximate the Green's functions on predefined interpolation points. Generally the number of interpolation points (equivalent to the rank) is much smaller than the number of unknowns in the coupling groups. The compressed Green's functions matrices are inserted in to the integral equations for acceleration. Linear computational complexity can be achieved in the low and medium frequency regime [57]–[60]. The second method is to factorize the impedance matrix directly, introducing dominant basis functions denoted as skeletons [61]–[65] and equivalent basis functions [32], [66]–[67] to represent the low-rank impedance matrix in a recursive formulation. $O(N)$ and $O(N \log N)$ computational complexity is achieved for low and high frequency problems, respectively.

An additional advantage of the low-rank methods is that they allow for efficient direct (non-iterative) solution and high efficiency preconditioning techniques for the matrix linear system [22], [68]–[71]. Avoiding the need for an iterative solver has several advantages. To name the most important ones, there is no need for preconditioning, which is often the bottle-neck of iterative methods, and the system can be solved for several excitation vectors simultaneously, for example in monostatic RCS computations. Most of the compressed direct MoM solvers proposed in the literature use either single-level [68] or nested LU factorization. An alternative approach, proposed in [69], uses a nested block-decomposition based on the Sherman–Morrison–Woodbury formula for the inverse of a partitioned matrix. In [71], this approach, known as multiscale compressed block decomposition (MSCBD), is shown to be considerably more efficient than nested LU decomposition, although both approaches have a high-frequency computational complexity of $O(N^2)$for the factorization step. Similarly, the H-Matrices [72], [73], H-LU [74]–[78], and H^2 matrix [79]–[83] methods are proposed for dense and sparse matrix inversion, respectively. Unfortunately, to date this approach has not yet been successfully applied to large-scale high frequency problems [82]. Recently, a new and promising approach has been proposed, using randomized compression algorithms to reduce the complexity of the butterfly construction [29]. The approach is shown to achieve the theoretical complexity of $O(N\log^2 N)$. The paper additionally presents an $O(N^{3/2}\log^2 N)$ direct solver, also based on randomized algorithms. More recently, the hierarchically off-diagonal low rank method has been proposed, where the dense matrix is decomposed as the product of several block diagonal matrices, direct inversion or highly efficient preconditioners can be constructed for multiscale simulations [87]–[91].

The remainder of the paper is organized as follows: in Section II, the challenges from multiscale simulations are proposed, the traditional low-rank methods are illustrated in Section III, the improved algorithm for ACA is proposed in Section IV, post compression low-rank methods are illustrated in Section V, the nested low-rank methods are illustrated in Section VI, and the direct inversion of low-rank methods is illustrated in Section VII. Finally, a brief conclusion is given in Section VIII.

II. CHALLENGES FROM MULTISCALE SIMULATIONS

The multiscale problems are typical mixed low and high frequency problems, where coexistence of dense meshes are

to capture the geometric details and of large-scale interactions. The difficulties associated to this scale variability are enhanced in analyses requiring a large frequency range, often with the requirement to change as little as possible the mesh over the frequency range of interest. The challenges for multiscale fast solvers are mainly:

(1) For standard fast multipole method, it would meet low-frequency break down when the group size is smaller than 0.2 wavelength. Several low frequency stable technologies have been developed to solve this problem [33]–[40].

(2) Ill-conditioned linear system, for dense mesh with small electrical size, the vector potential is much smaller than the scalar potential in the conventional electric field integral equation (EFIE). Since the dominating scalar potential matrix is singular, the matrix system is ill-conditioned and even breaks down due to finite machine precision. The impedance matrix tends to be increasingly ill-conditioned as the increase of ratio between the maximum and minimum discretization size.

The so-called low-frequency breakdown is an important aspect of the multiscale simulations; this problem arises in two different ways, that are not separated in practice. On the one hand, low-frequency issues are associated to dense-meshes; by itself this is independent of the employed specific fast factorization approach; it is associated to the problem conditioning, and as such mitigated or solved by conditioning techniques. However, this is crucial in association with fast factorizations: these are always approximations, and hence introduce a perturbation to the (ideal) system matrix, whose effect on the solution is crucially dependent on matrix conditioning. It is interesting to note that this impacts direct inversions [22], [68]–[71], as well as the more standard iterative solvers.

The other low-frequency issue is associated to possible breakdown of the fast factorization itself, as well documented for the MLFMA [14], [15]. In low-rank factorizations, the computation accuracy is independent of the group size, they are numerically stable at low frequency. In the algorithm, the number of levels for geometric or matrix decomposition is determined by the average number of unknowns at leaf level. The low-rank factorization is error controllable by predetermined threshold [23], [24] and number of equivalence points [27], [66], [67].

The hierarchical preconditioner [19] and direct solvers [68]–[71] are proposed with low-rank factorization methods to solve the matrix system from ill-conditioned matrix system.

III. TRADITIONAL LOW-RANK METHODS

When the object is subdivided into groups according to the average number of unknowns in the clustered groups, the MoM discretization of the integral equation yields the linear system

$$\mathbf{ZI} = (\mathbf{Z}_{\text{near}} + \mathbf{Z}_{\text{far}})\mathbf{I} = \mathbf{V}, \tag{1}$$

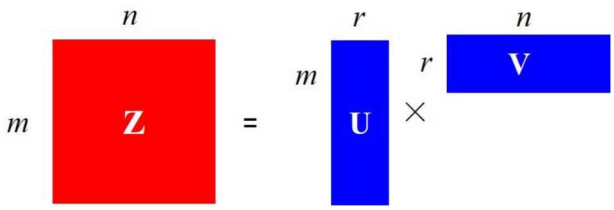

FIGURE 1. Impedance matrix subblock coupling two far groups is approximated with traditional low-rank factorization methods.

where $\mathbf{Z}_{\text{near}}$ denotes the near field coupling evaluated with MoM directly, and $\mathbf{Z}_{\text{far}}$ denotes the far field couplings, which can be evaluated with low-rank matrix factorization method in this review. The admission condition for the two far couplings groups t and s is defined as

$$R(s, t) \geq 2D_l, \tag{2}$$

where $R(s, t)$ is the center-to-center distance between groups t and s, D_l is the group size at the level l. As shown in Fig. 1, for traditional low-rank methods, the impedance matrix subblock between two far field groups can be approximated as

$$\mathbf{Z}_{m\times n} = \mathbf{U}_{m\times r} \times \mathbf{V}_{m\times r}. \tag{3}$$

The multiplication of $\mathbf{U}$ and $\mathbf{V}$ gives a good approximation of the actual interaction matrix with predetermined threshold. The memory storage of the interaction matrix will be $O(r(m + n))$ instead of $O(mn)$, and the matrix multiplication with a vector $\mathbf{I}$ of size n requires only $O(r(m + n))$ operations instead of $O(mn)$. The differences between traditional low-rank factorization methods are the strategies to construct the low-rank approximation matrices in Eq. (3). Equivalent sources are introduced in MDA to obtain the low rank approximation, where the rank is equal to the number of equivalent points [21]. The dominant columns and rows are selected adaptively in ACA [22]–[24] with a predefined precision. Uniform or random sampling of the columns and rows are proposed in the UV method [24], [25]. Further combination of the UV method with a multiresolution preconditioner is proposed for the simulation of frequency selective surface (FSS) arrays in [26].

Since the low-rank decompositions in Eq. (3) are implemented for each pair of two groups, even for a multilevel version, the low-rank approximation matrices should be constructed repeatedly, which in turn leads to high computational and memory resource costs. As a result, they are only applicable for electrical small or medium problems. The computational complexity in this regime is $O(N^{4/3} \log N)$ [24].

IV. IMPROVED ACCURACY FOR THE ADAPTIVE CROSS APPROXIMATION ALGORITHM

The most popular algorithm for low-rank compression of off-diagonal impedance matrix blocks is the ACA algorithm. The ACA efficiently produces a low-rank approximation of

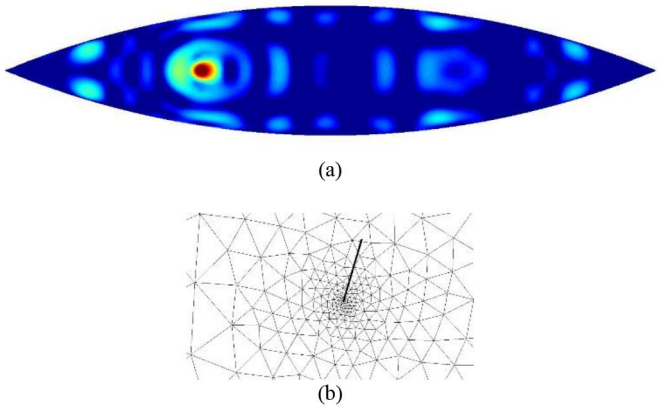

FIGURE 2. (a) Monopole mounted on a PEC Ogive, surface current distribution at the fundamental resonance, full non-compressed solution, (b) refined triangulation of the Ogive surface near the monopole base.

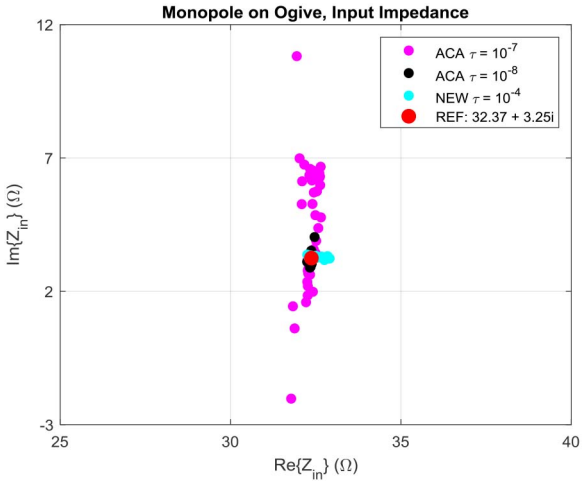

FIGURE 3. Input impedance of a thin $\lambda/4$-monopole mounted on a PEC Ogive. Computed with conventional ACA and with the new adapted ACA proposed in [94]. Each computation was repeated 50 times, with randomly chosen ACA initial row indices. The reference value was obtained by full LU decomposition of the impedance matrix.

a matrix with a relative error that is governed by a preset accuracy threshold value. A number of publications have drawn the attention to the fact that sometimes, the underlying mechanism fails and the error is much larger than expected [92], [93]. This phenomenon is particularly relevant for direct solution of large multiscale problems mainly for two reasons: Firstly, direct solution methods are most efficient when all off-diagonal blocks are compressed, including those that represent touching regions. This is sometimes referred to as the "weak admissibility criterion". Any off-diagonal blocks that are not compressed will lead to "fill-in" of uncompressed blocks in the direct inversion stage and rapidly growing computational and memory costs. However, the error in the solution is very sensitive to compression errors in these "near field" blocks. Secondly, multiscale problems typically lead to badly conditioned impedance matrices, due to the disbalance in the matrix element sizes. While direct methods generally deal with badly conditioned matrices better than iterative methods, compression errors will have a much larger effect on the error in the solution if the matrix is badly conditioned.

Recently, a thorough study of the above-mentioned phenomenon has been published in [94]. This paper proposes a modified algorithm that largely remedies the problem. The paper demonstrates that there exists an inherent random uncertainty in the conventional ACA algorithm; the choice of the row-index that initializes the ACA iterative process, which highly influences the approximation error. The proposed adaptation in [94] makes this randomness explicit and allows to control its effect. We refer to [94] for the details of the proposed algorithm. In [94], the improved accuracy is illustrated with a numerical example of an iterative solver accelerated with low-rank compression and a "strong" admissibility criterion. Here we present an example demonstrating that the improvement is much more important for multiscale problems that are solved by direct inversion with weak admissibility.

The target under investigation is the PEC ogive from [95] (major axis 25.4 cm), shown in Fig. 2 (a), but with a very small (0.0254×5.08 mm) flat strip monopole antenna mounted on the ogive surface. The mesh-size of the antenna and surrounding region on the ogive surface is much smaller than that of the rest of the ogive, as illustrated in Fig. 2 (b). The entire structure is meshed into 19,692 triangular facets on which 29,338 RWG basis functions are defined. The excitation is a 14.4 GHz delta-gap on the monopole base, near the first resonance. The problem is solved using the EFIE formulation, initially by full non-compressed LU decomposition to obtain a reference solution. Subsequently the structure is decomposed according to a 6-level binary tree and ACA compression is applied to all off-diagonal blocks. The compressed system is then solved by direct inversion using the MSCBD algorithm [71].

We solve the system repeatedly, with randomly chosen initial row index for all ACA compressed blocks. Using a typical ACA accuracy threshold of $\tau = 10^{-3}$ or $\tau = 10^{-4}$, the solution fails completely, as a consequence of the high condition number of the matrix, due to the presence of highly unequal matrix elements. As seen in Fig. 3, with $\tau = 10^{-7}$ the error, defined as the difference in the monopole input impedance with that of the full uncompressed solution, is still considerable for a large proportion of the simulations. Only for $\tau = 10^{-8}$ the result stabilizes. In contrast, when we implement the adaptation of [94], a stable solution is obtained already with $\tau = 10^{-4}$. Essentially, the algorithm automatically detects that the convergence of the ACA fails for touching regions involving highly disbalanced matrix elements and determines the necessary threshold for these cases. The improvement in memory requirements and computation time are considerable, as evidenced by Table 2 (All computations were done in MATLAB, double precision, on a laptop with 16 GB of RAM and an Intel Core i7-8565U CPU at 1.80GHz).

TABLE 2. Average matrix sizes and computation times for ACA compressed MoM analysis of monopole mounted on PEC Ogive.

	Matrix setup time (s)	Matrix size (MB)	Factorization time (s)	Factorization size (MB)	Solution time (s)
[24] ACA $\tau = 10^{-7}$	272	1405	219	1424	1.4
[24] ACA $\tau = 10^{-8}$	419	1723	324	1727	1.6
[94] NEW $\tau = 10^{-4}$	98	709	78	869	1.1

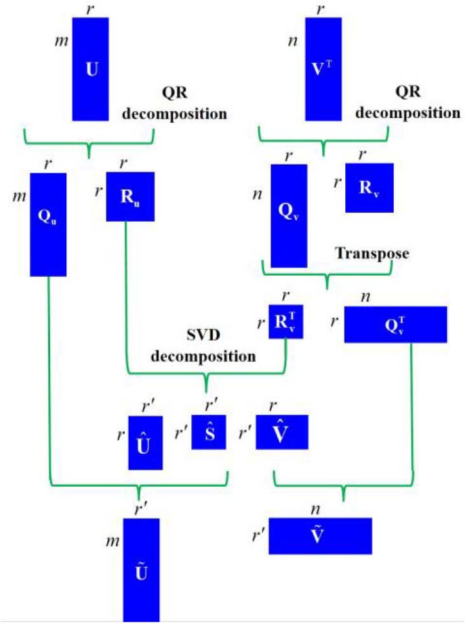

FIGURE 4. SVD post compression of typical low-rank methods.

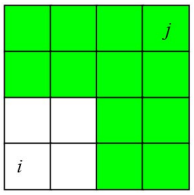

FIGURE 5. Peer-level far field coupling groups (green color region) of group *i*.

V. POST COMPRESSION LOW-RANK METHODS

Singular value decomposition (SVD) is usually employed for the low-rank approximation matrices post compression [27]. The post compression process for typical low-rank methods is shown in Fig. 4. Eq. (3) can be rewritten as

$$\mathbf{Z}_{m\times n} = \mathbf{U}_{m\times r} \times \mathbf{V}_{r\times n} = \tilde{\mathbf{U}}_{m\times r'} \times \tilde{\mathbf{V}}_{r'\times n}, \tag{4}$$

where r' is smaller than r.

However, as mentioned in Section III, for each pair of far coupling groups, the low-rank approximation matrices $\mathbf{U}$ and $\mathbf{V}$ need to be computed again, which is much less efficient than the fast multipole method (FMM) [14], [15], in which the aggregation and disaggregation operators of a group are defined only once regardless of the far field interaction groups. For each group, the multilevel matrix compression method (MLMCM) [30] and reciprocal MLMCM (rMLMCM) [51] are proposed respectively to formulate a single aggregation and disaggregation operator of the far field interaction groups.

Fig. 5 shows the peer-level far field coupling lists of group i. Let $[\mathbf{Z}_{ij}]_{m\times n}$ represent the impedance submatrix between two groups i and j. The MLMCM starts by collecting the columns $\mathbf{Z}_{i,1}, \mathbf{Z}_{i,2}, \mathbf{Z}_{i,3}\cdots$ of $[\mathbf{Z}_{ij}]_{m\times n}$ of the peer-level far field groups of group i, and then a modified Gram-Schmidt (MGS) algorithm is applied and the *receiving compression matrix* is defined as

$$\mathbf{U}_i = \mathrm{MGS}\big[\mathbf{Z}_{i,1}, \mathbf{Z}_{i,2}, \mathbf{Z}_{i,3}\cdots\big]. \tag{5}$$

Then the coupling matrix $\mathbf{Z}_{m\times n}$ is factorized as:

$$\big[\mathbf{Z}_{ij}\big]_{m\times n} = [\mathbf{U}_i]_{m\times r}\big[\mathbf{D}_{ij}\big]_{r\times r}\big[\mathbf{V}_j\big]_{r\times n}. \tag{6}$$

where r is the ε-rank of $\mathbf{Z}_{ij}$, and m and n are the number of basis functions in groups i and j, respectively. $\mathbf{U}_i$, $\mathbf{D}_{ij}$, and $\mathbf{V}_j$ are three relatively small dense matrices. Inspired by [31] and [96], [97], a reciprocal algorithm rMLMCM [51] is developed over [30], where the radiation compression matrix $\mathbf{V}_j$ satisfies

$$\mathbf{V}_j = \mathbf{U}_j^{\mathrm{T}}. \tag{7}$$

To compute all the column vectors in Eq. (5) is time consuming, so a subset of the column vectors (the skeletons) is sampled with ACA, yielding an error controllable procedure. By construction, $\mathbf{U}_i$ satisfies $\mathbf{U}_t^{\dagger}\mathbf{U}_t = \mathbf{I}$, with $\mathbf{I}$ the identity matrix, and $(\bullet)^{\dagger}$ denotes conjugate transpose; as a consequence, the *translation matrix* $\mathbf{D}_{ts}$ between groups t and s can be explicitly written as:

$$\mathbf{D}_{ij} = \mathbf{U}_i^{\dagger}\,\mathbf{U}_i\,\mathbf{D}_{ij}\,\mathbf{V}_j\,\mathbf{V}_j^{\dagger} = \mathbf{U}_i^{\dagger}\,\mathbf{Z}_{ij}\,\mathbf{V}_j^{\dagger}. \tag{8}$$

Note that $\mathbf{Z}_{ij}$ in (8) is not explicitly computed, it is evaluated with ACA instead. The translation matrix $[\mathbf{D}_{ij}]_{r\times r}$ has strongly reduced dimensions with respect to the original matrix $[\mathbf{Z}_{ij}]_{m\times n}$.

The computational complexity of the rMLMCM is $O(NlogN)$ for medium electrical sizes [51]. The computational complexity is tested for a series of cubic cavities as shown in the inset in Fig. 6 (a). The edge length of the cube is 2m with a fixed average mesh size of 6e-2λ. The simulated frequencies are 125MHz, 250MHz, 500MHz, and 1GHz, corresponding to a total number of unknowns equal to 2,644, 10,495, 42,080, and 168,520, respectively. Fig. 6 (a) and (b) show the computational complexity of a matrix-vector product (MVP) and memory requirements of the solver [51].

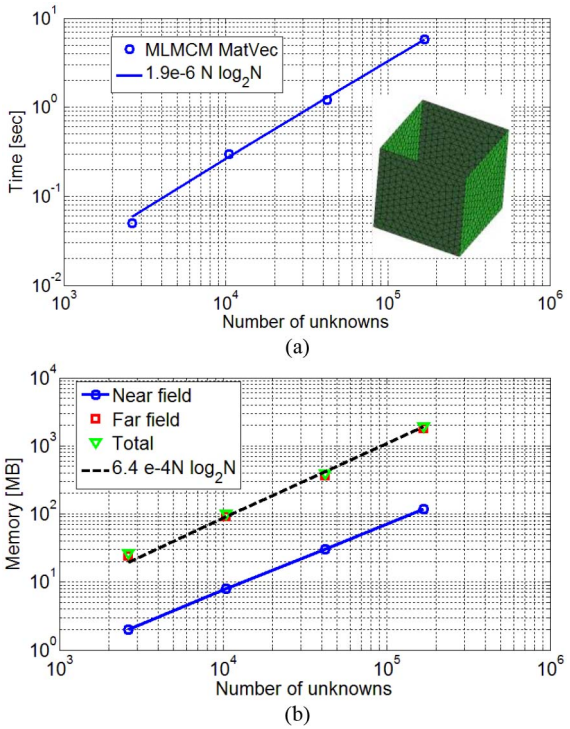

FIGURE 6. Computational complexity of the proposed doubly hierarchical MoM (a) MVP time (b) storage requirements. The figures are reprinted from [51].

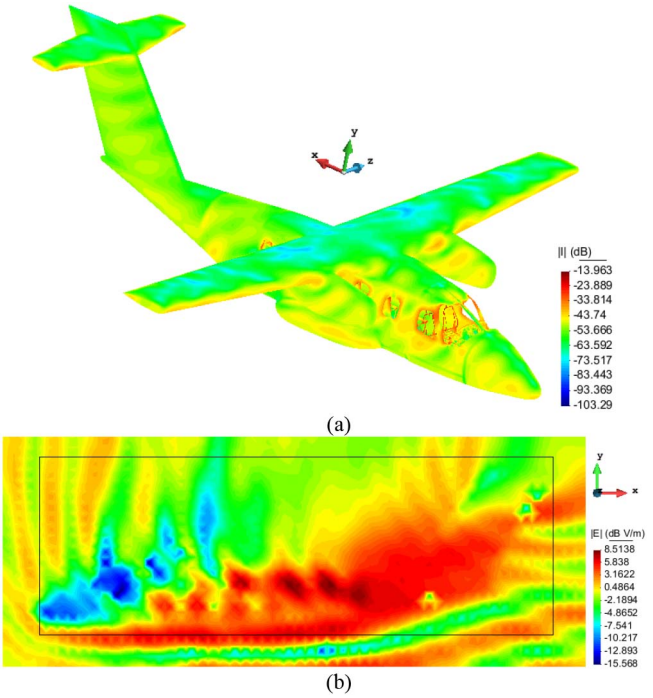

FIGURE 7. Morphed EV55 aircraft simulated at 244 MHz: (a) surface current density; (b) Electric field on a symmetry plane of the EV55 aircraft. The figures are reprinted from [51].

For realistic large and multiscale problems, two strategies are introduced for the rMLMCM: Firstly, a hybrid rMLMCM/MLFMA [51] is proposed, where the rMLMCM is employed to evaluate the low and mid frequency regime due to the dense mesh of the fine structures installed on the large platforms. The MLFMA [14], [15] is employed to evaluate the high frequency regime, where the rMLMCM will suffer computation performance degeneration. Secondly, the hierarchical multiresolution (MR) preconditioners [98], [99] are introduced to rMLMCM/MLFMA to relieve the convergence challenges stemming from the ill-conditioned matrix equation due to multiscale discretization.

A morphed model of a realistic Evektor EV55[1] aircraft is simulated, where all internal details, such as passenger seats and the instrument board are considered. The aircraft is 14.2m long, the wingspan is 16.1m, corresponding to 11.5 and 13λat 224MHz, respectively. The mesh size ranges from 3.6e-3λ to 6.3e-2λ, leading to 171,763 unknowns. The aircraft is illuminated by a plane wave impinging from $(\theta^i = 90°, \varphi^i = 225°)$. Single-level rMLMCM and 5-level MLFMA are employed, corresponding to a near field evaluation region of 0.1λ.

With the MR-ILU precondtioner, the time and memory requirements are 625 MB and 3.9 minutes, while for the typical ILU preconditioner, 4.1 hours and 14.1 GB are required. However, the number of iterations required is comparable (the iterations of MR-ILU is 810, and the ILU is 906). Hence, significant savings in total simulation time and memory requirements are obtained. Meanwhile, better convergence performances are achieved with the MR-ILU preconditioner [19], [99]. Fig. 7 (a) and (b) show the surface current density and electric field distribution inside the aircraft in the x-y plane of the morphed EV55 aircraft, which demonstrates the capability of the proposed rMLMCM/MLFMA to model realistic high definition multiscale problems [51].

1. http://www.evektoraircraft.com/en/aircraft/ev-55-outback/overview

VI. NESTED LOW-RANK METHODS

As explained in Section V, the approximation framework of MLMCM and rMLMCM is similar to the FMM. It is well known that the MLFMA achieves lower computational complexity due to the nested expression denoted as aggregation and disaggregation between neighboring levels. Can we construct a nested approximation framework like the MLFMA? The low-rank approximation matrices are only computed at the leaf level, the higher level matrices are expressed by those at leaf level. This leads to the nested equivalent source approximation (NESA) for low-to-high frequency, proposed in [66], [67], [100], and [101].

A. NESTED MATRIX COMPRESSION VIA EQUIVALENT AND SKELETON BASES

To obtain the nested matrix compression [32], the skeleton and equivalent bases are introduced respectively in interpolation decomposition (ID) [61], [62] and nested equivalent source approximation (NESA) [66], [67] methods. As shown

in Fig. 8 (a), the inner equivalent surface $\sum_{\tau}^{s}$ and outer testing surface $\sum_{\sigma}^{s}$ are defined respectively for group i , the radius of the inner and outer sphere surface are $S/2$ and $3S/2$ as in Fig. 8 (c), S is the group size.

When considering the coupling between group s and o, our goal is to find the coefficients of the equivalent source $\boldsymbol{\tau}_s$ on $\sum_{\tau}^{s}$, which radiates the same field on the testing surface $\sum_{\sigma}^{s}$ as the actual source in group s. Then the far coupling with group s can be substituted by the coupling with the equivalent source $\boldsymbol{\tau}_s$. The coefficients can be obtained by equating the fields radiated by two sets of equivalent source $\boldsymbol{\tau}_s$ and actual source on the testing basis functions $\boldsymbol{\sigma}_s$ in a weak sense

$$\mathbf{Z}_{\sigma_s,s}\mathbf{I}_s = \mathbf{Z}_{\sigma_s,\tau_s}\mathbf{I}_{\tau_s}, \tag{9}$$

where $\mathbf{I}_s$ and $\mathbf{I}_{\tau_s}$ are the current coefficients of the actual and equivalent basis functions, respectively. Matrix $\mathbf{Z}_{i,j}$ in eq. (9) and following is defined as the coupling impedance matrix between the basis i and j. Then equivalent current coefficients $\mathbf{I}_{\tau_s}$ can be derived as:

$$\mathbf{I}_{\tau_s} = \left(\mathbf{Z}_{\sigma_s,\tau_s}\right)^{\dagger}\mathbf{Z}_{\sigma_s,s}\mathbf{I}_s. \tag{10}$$

Here $.^{\dagger}$ denotes the pseudo-inverse of the matrices formed by basis functions on the equivalent and testing surface. We define the *radiation matrix* for group s as

$$\mathbf{V}_s = \left(\mathbf{Z}_{\sigma_s,\tau_s}\right)^{\dagger}\mathbf{Z}_{\sigma_s,s}. \tag{11}$$

Similarly, when considering the observation group o, the *receiving matrix* can be derived as

$$\mathbf{U}_o = \mathbf{Z}_{o,\sigma_o}\left(\mathbf{Z}_{\tau_o,\sigma_o}\right)^{\dagger}, \tag{12}$$

$\boldsymbol{\tau}_o$ and $\boldsymbol{\sigma}_o$ are the equivalent and testing basis functions of group o. The *translation matrix* between group s and o is defined as

$$\mathbf{D}_{o,s} = \mathbf{Z}_{\tau_o,\tau_s}. \tag{13}$$

Then the far coupling between group s and o can be evaluated with NESA as

$$\mathbf{Z}_{o,s} = \mathbf{U}_o\mathbf{D}_{o,s}\mathbf{V}_s. \tag{14}$$

It can be found from Eq. (14), when coupling with the far group lists at the peer level, only one radiation and receiving matrix needs to be computed.

The difference between ID [61], [62] and NESA [66], [67] is the equivalent process as in Fig. 8 (a) and (b). For ID, only the testing equivalent surface $\sum_{\sigma}^{s}$ is constructed, the skeleton basis functions (denoted as red arrows) are sampled from the actual basis function in group s instead of the equivalent basis functions. For ID, combined with ACA, the number of skeleton basis functions at low levels is usually smaller than the number of equivalent basis functions of NESA. As a result, a faster matrix-vector product is achieved. However, with NESA, the equivalent and testing basis functions are the same for every group hence the translation invariant symmetry can be employed to reduce the setup low-rank approximation time and memory requirements significantly [66], [67].

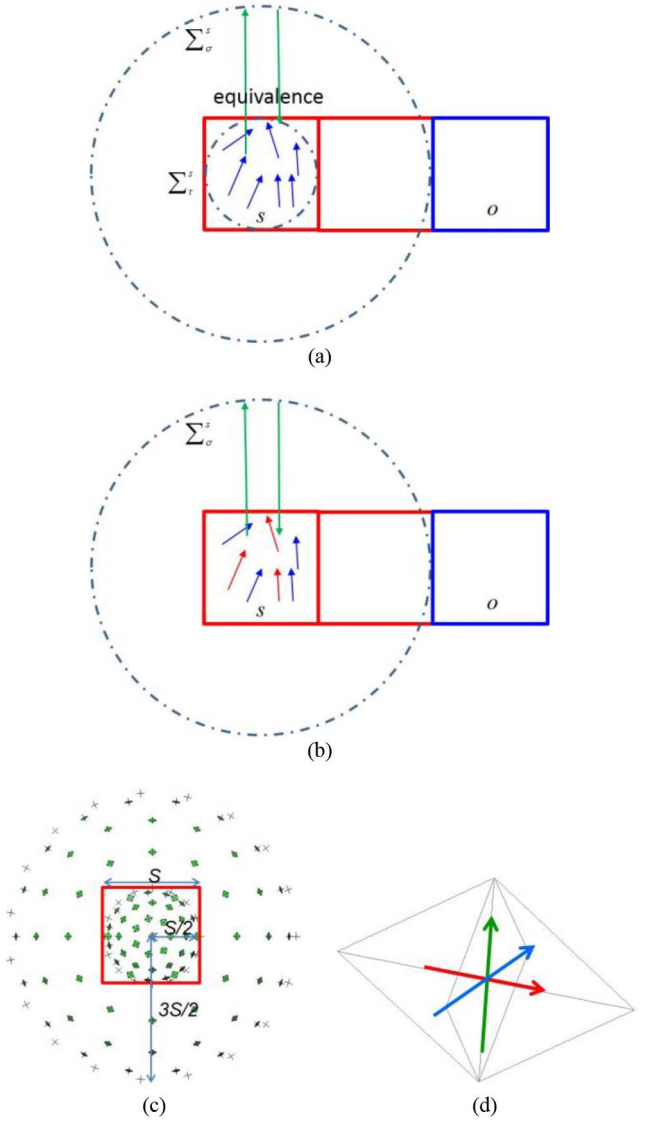

FIGURE 8. Nested matrix approximation for two far coupling groups *i* and *j* via equivalent and dominant bases, (a) equivalent bases on the inner sphere surface are obtained by enforcing testing on the outer sphere surface, (b) dominant (skeleton) bases are obtained from the original basis functions by enforcing testing on the outer sphere surface, (c) three dimensional (3D) view of the inner equivalent surface and outer testing surface, at each sampling points on the sphere surface, (d) three orthogonal basis functions are defined, respectively.

A mixed-form skeleton and NESA algorithm is proposed in [101] for multiscale simulations, where at the low-levels, the ID is employed to achieve smaller ranks, at the high-levels, the NESA is employed to achieve setup memory and time savings. Three orthogonal RWG basis functions [2] are defined as the equivalent and testing basis functions for NESA as in Fig. 8 (d). In the implementation, the size of the triangles is smaller than $\lambda/30$ to achieve high computation precision with only one integration point on the equivalent testing surface when evaluating the impedance matrix in Eq. (9) to Eq. (13).

The low-rank approximation for NESA at leaf level is obtained in Eq. (14). Like in the MLFMA, the *transfer matrices* [66], [67] between neighboring levels are defined to obtain a nested approximation

$$\mathbf{V}_{s^p}^{l-1} = \mathbf{C}^{l-1,l}\mathbf{V}_s^l \tag{15}$$

$$\mathbf{U}_{o^p}^{l-1} = \mathbf{U}_o^l\mathbf{B}^{l,l-1}. \tag{16}$$

The low-rank approximation matrices $\mathbf{V}_{s^p}^{l-1}$ and $\mathbf{U}_{o^p}^{l-1}$ of the parent groups s^p and o^p at level $l-1$ can be expressed with the low-rank approximation matrices of child groups at level l via the transfer matrices $\mathbf{C}^{l-1,l}$ and $\mathbf{B}^{l,l-1}$. With the nested low-rank approximation linear and $O(N\log N)$ complexity can be achieved for low and high frequency multiscale simulations, respectively. As demonstrated in [66], [67], the low-rank approximation is error controllable through the number of equivalent points Q on the equivalent sphere surface.

B. LOW-TO-HIGH FREQUENCY NESTED EQUIVALENT SOURCE APPROXIMATIONS

As we know, the key point which affects the computational complexity is the average rank. In the low frequency regime, the average rank remains constant while in the high frequency regime, the average rank increases very fast with increasing group size, which leads to a high complexity, approaching that of the full MoM [24], [49].

In the nested approximation, the admission conditions are defined differently for low and high frequency regime. $D_0 = \lambda$ is the threshold group size between low-frequency and high-frequency couplings. In the low frequency regime ($D_l < D_0$), the admission condition is the same as traditional low-rank methods, i.e., groups s and o are not neighbors

$$R(s,o) \geq 2D_l. \tag{17}$$

where $R(s,o)$ is the center-to-center distance between groups s and o. Nested approximation as shown in Fig. 8 can be employed directly [66]. Conversely, in the high frequency regime ($D_l \geq D_0$), the directional low-rank property is exploited to guarantee constant rank in certain directions [102], [103]. As shown in Fig. 9, the directional low-rank property is invoked to define pyramids spanning an angle of $O(\lambda/D_l)$. The peer level far coupling region of groups s and o at level l is defined as

$$\frac{R(s,o)}{\lambda} \geq \left(\frac{D_l}{\lambda}\right)^2 \tag{18a}$$

$$\frac{R(s_p,o_p)}{\lambda} < \left(\frac{D_{l-1}}{\lambda}\right)^2, \tag{18b}$$

where $D_{l-1} = 2D_l$ is the parent group size at level $l-1$, i.e., the far coupling interaction list of a source group s includes groups o satisfying the admissibility condition (18a), subject to their parents s_p and t_p not satisfying (18b). The directional admission conditions shown in Fig. 9 guarantee the "hierarchical directions," i.e., each direction of a group is completely enclosed by the directions of its child groups [102], [103]. This in turn guarantees that, if two groups satisfy the admissibility condition of (18a), then their children satisfy the admissibility condition naturally. The wideband NESA automatically switching between low and high frequency can be achieved with the admission conditions defined in Eq. (17) and (18) [67], [101].

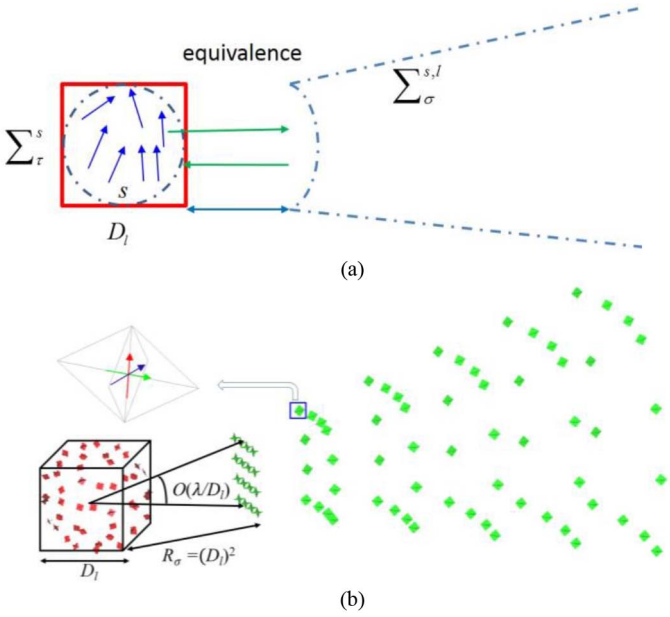

FIGURE 9. Directional admission conditions for the NESA at high frequency regime, (a) 2D view, (d) 3D view, the nested approximation of Eq. (9)-Eq. (10) is obtained with the inner equivalent sphere surface and outer testing pyramid, The angle of the pyramid is $O(\lambda/D_l)$ and the distance from the equivalent sphere to the pyramid is $(D_l)^2$, $D_l$is the group size at level l.

TABLE 3. Computation complexity of the existing low-rank methods.

Method	Computational complexity	
	Medium size	Large size
ACA	$O(N^{4/3}\log N)$	—
MLMCM	$O(N\log N)$	—
NESA	$O(N)$	$O(N\log N)$

To achieve the $O(N\log N)$ computation complexity of the MLFMA in the high frequency regime, two new contributions are introduced in NESA: firstly, the nested low-rank approximation in Section VI-A, secondly, the rank Q is independent of the group size [67], exploiting the directional low-rank property [102], [103]. Even more important, the introduced equivalent and testing surfaces lead to an intrinsically multiscale family of auxiliary sources, improving field representation in multiscale problems, which in turn leads to a significant improvement in convergence performance [66], [67]. Table 3 summarizes the computational complexity of the low-rank factorization methods, it can be found the NESA can achieve much better performance both in low frequency and high frequency regime.

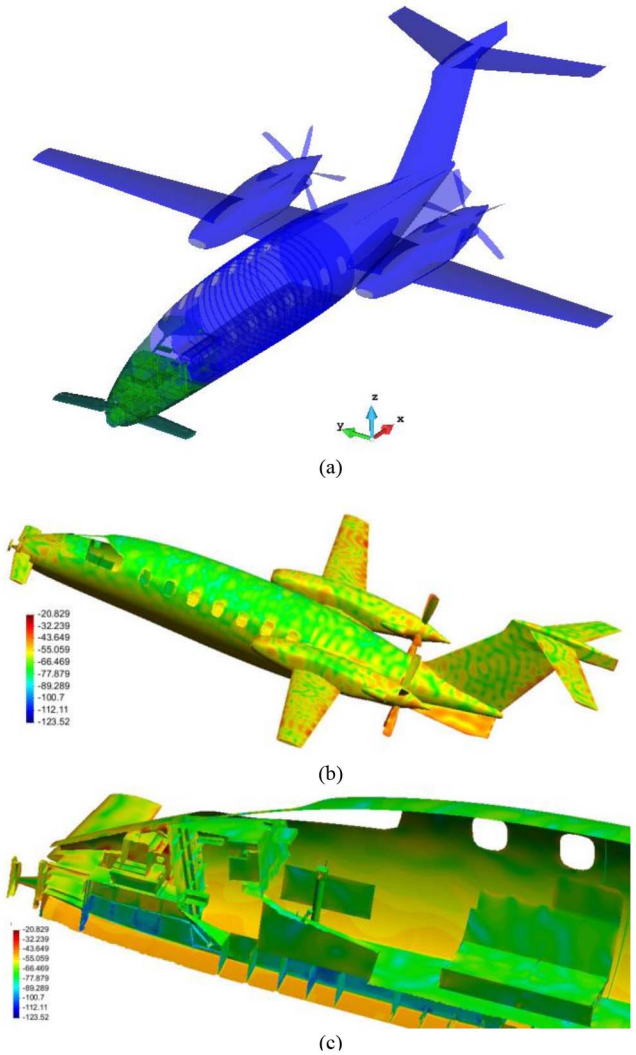

FIGURE 10. Simulated morphed P180 aircraft model at 686 MHz. A flexible GMRES iterative solution with 10 inner iterations is employed, 100 iterations are required to achieve convergence to a residual of 1e-3, (a) mesh model, surface currents of the aircraft (b) and inside (c). The figures are reprinted from [67].

C. NUMERICAL RESULTS FOR MULTISCALE PROBLEMS

The NESA is employed in the European Community's Seventh Framework Programme FP7/2007-2013 HIRF SE Project for realistic multiscale simulation from low to high frequency [66], [67].

A morphed P180[2] aircraft shown in Fig. 10 (a), has been analyzed. The aircraft is 12.1 m long, and its wingspan is 13.8 m, corresponding to, respectively, 27.6 and $31 \cdot 5\lambda$ at 686 MHz. All internal details such as passenger seats, antenna array, and the instrument board are considered in the model. The aircraft is illuminated by a plane wave impinging along $\hat{y}$ directions with the electric field polarized along $\hat{z}$ as in Fig. 10 (a). The number of unknowns is 1,086,083 with discretization h/λ ranging from 2.3e-3 to 8.0e-2; two-levels of low frequency and three-levels of high frequency are defined, and the MR-ILU preconditioner [98], [99] is used to improve the convergence. Factorization time and memory required are 1.8 h and 9.1 GB; The MVP time is 28 s, and overall solution time of the matrix equation amounts to 7.8 h. Fig. 10 (b) and (c) show the surface currents of the details of the aircraft and the details inside, respectively.

2. http://www.piaggioaero.com/#/en/products/p180-avanti-ii/overview

VII. FAST DIRECT INVERSION LOW-RANK METHODS

A. H-MATRICES METHOD

When we have the low-rank factorization of the impedance matrix, it can be represented by the H-matrices format [22]. The inverse matrix can also be expressed accordingly. A general dense matrix $\mathbf{Z}$ from the MoM at level $l = 1$ can be written as

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_{11}^{(1)} & \mathbf{Z}_{12}^{(1)} \\ \mathbf{Z}_{21}^{(1)} & \mathbf{Z}_{22}^{(1)} \end{pmatrix}, \tag{19}$$

where $\mathbf{Z}_{ij}^{(l)}$ denotes the coupling matrix between group i and j at level l. The off-diagonal matrices $\mathbf{Z}_{12}^{(1)}$ and $\mathbf{Z}_{21}^{(1)}$ can be evaluated with the low-rank matrix method, while (19) can be used recursively to partition the diagonal matrix $\mathbf{Z}_{11}^{(1)}$ and $\mathbf{Z}_{22}^{(1)}$. The off-diagonal matrices at level $l = 2$ can be characterized with the low-rank matrix method as

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_{11}^{(1)} & \mathbf{U}_{12}^{(1)}\mathbf{V}_{12}^{(1)} \\ \mathbf{U}_{21}^{(1)}\mathbf{V}_{21}^{(1)} & \mathbf{Z}_{22}^{(1)} \end{pmatrix}, \tag{20}$$

where

$$\mathbf{Z}_{11}^{(1)} = \begin{pmatrix} \mathbf{Z}_{11}^{(2)} & \mathbf{U}_{12}^{(2)}\mathbf{V}_{12}^{(2)} \\ \mathbf{U}_{21}^{(2)}\mathbf{V}_{21}^{(2)} & \mathbf{Z}_{22}^{(2)} \end{pmatrix} \tag{21a}$$

$$\mathbf{Z}_{22}^{(1)} = \begin{pmatrix} \mathbf{Z}_{33}^{(2)} & \mathbf{U}_{34}^{(2)}\mathbf{V}_{34}^{(2)} \\ \mathbf{U}_{43}^{(2)}\mathbf{V}_{43}^{(2)} & \mathbf{Z}_{44}^{(2)} \end{pmatrix}. \tag{21b}$$

The inversion of (19) can be obtained with the H-matrices method [22]

$$\mathbf{Z}^{-1} = \begin{pmatrix} \left(\mathbf{Z}_{11}^{(1)}\right)^{-1} + \left(\mathbf{Z}_{11}^{(1)}\right)^{-1}\mathbf{Z}_{12}^{(1)}\mathbf{S}^{-1}\mathbf{Z}_{21}^{(1)}\left(\mathbf{Z}_{11}^{(1)}\right)^{-1} & -\left(\mathbf{Z}_{11}^{(1)}\right)^{-1}\mathbf{Z}_{12}^{(1)}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{Z}_{21}^{(1)}\left(\mathbf{Z}_{11}^{(1)}\right)^{-1} & \mathbf{S}^{-1} \end{pmatrix}, \tag{22}$$

with the Schur complement $\mathbf{S} = (\mathbf{Z}_{22}^{(1)}) - \mathbf{Z}_{21}^{(1)}(\mathbf{Z}_{11}^{(1)})^{-1}\mathbf{Z}_{12}^{(1)}$, the low computation complexity matrix-matrix addition and multiplication can be achieved by the H-matrices. The H-matrices is widely applied to the MoM and finite element method (FEM) for dynamic electromagnetic problems [73]–[83].

Specially, the H-matrices method is further extended to hierarchically semiseparable (HSS) [84]–[86] and hierarchical off-diagonal low-rank systems (HODLR) [87]–[91] algorithms for different admissibility conditions. For dynamic electromagnetic problems, the sub-impedance matrix blocks produced by neighboring couplings groups in MoM are almost full-rank, as a result, the HODLR algorithm are widely employed for large multiscale simulations [87]–[91].

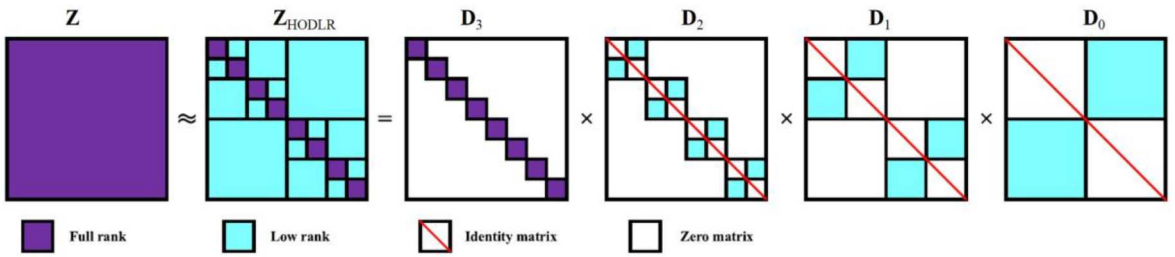

FIGURE 11. HODLR matrix and factorization with level $L = 3$. Z_{HODLR} denotes the hierarchically data-sparse representation of the impedance matrix Z. D_3 is the block-diagonal matrix, D_2 to D_0 is the remaining matrix after extracting the diagonal block matrix. The inversion of block-diagonal matrix can be obtained directly, while the inversion of remaining matrix can be obtained by Sherman–Morrison–Woodbury algorithm [87]–[91]. The figures are reprinted from [91].

B. HIERARCHICALLY OFF-DIAGONAL LOW-RANK ALGORITHM

If we extract the diagonal matrix at each level of (20), the HODLR format can be achieved as the product of several diagonal matrices as shown in Fig. 11. The matrix equation in (21a) can be expressed with a two-level HODLR matrix format [87]–[91]

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_{11}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{22}^{(1)} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{I} & \tilde{\mathbf{U}}_{12}^{(1)}\mathbf{V}_{12}^{(1)} \\ \tilde{\mathbf{U}}_{21}^{(1)}\mathbf{V}_{21}^{(1)} & \mathbf{I} \end{pmatrix} = \mathbf{D}_1\mathbf{D}_0, \tag{23}$$

$\mathbf{I}$ is the identity matrix, the matrices $\tilde{\mathbf{U}}_{12}^{(1)}$ and $\tilde{\mathbf{U}}_{21}^{(1)}$ are updated by extracting the diagonal matrix $\mathbf{Z}_{11}^{(1)}$ and $\mathbf{Z}_{22}^{(1)}$

$$\tilde{\mathbf{U}}_{12}^{(1)} = \left(\mathbf{Z}_{11}^{(1)}\right)^{-1}\mathbf{U}_{12}^{(1)} \tag{24a}$$

$$\tilde{\mathbf{U}}_{21}^{(1)} = \left(\mathbf{Z}_{22}^{(1)}\right)^{-1}\mathbf{U}_{21}^{(1)}. \tag{24b}$$

Similarly, equation (23) can be further extended to an L level HODLR matrix as

$$\mathbf{Z}_{\text{HODLR}} = \mathbf{D}_L\mathbf{D}_{L-1}\cdots\mathbf{D}_3\mathbf{D}_2\mathbf{D}_1\mathbf{D}_0. \tag{25}$$

The inversion of the $\mathbf{Z}_{\text{HODLR}}$ can be easily obtained

$$\mathbf{Z}_{\text{HODLR}}^{-1} = \mathbf{D}_0^{-1}\mathbf{D}_1^{-1}\mathbf{D}_2^{-1}\cdots\mathbf{D}_{L-1}^{-1}\mathbf{D}_L^{-1}. \tag{26}$$

The inversion of $\mathbf{D}_L$ can be computed directly, while the inversion of the $\mathbf{D}_{L-1}$ to $\mathbf{D}_0$ can be computed with Sherman–Morrison–Woodbury algorithm, without loss of generality, D_0^{-1} can be computed as

$$\mathbf{D}_0^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} - \begin{pmatrix} \mathbf{0} & \tilde{\mathbf{U}}_{12}^{(1)} \\ \tilde{\mathbf{U}}_{21}^{(1)} & \mathbf{0} \end{pmatrix} \mathbf{S}_1^{-1} \begin{pmatrix} \mathbf{V}_{12}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{21}^{(1)} \end{pmatrix} \tag{27}$$

$$\mathbf{S}_1 = \begin{pmatrix} \mathbf{I} & \left(\mathbf{V}_{12}^{(1)}\right)^T\mathbf{U}_{12}^{(1)} \\ \left(\mathbf{V}_{21}^{(1)}\right)^T\mathbf{U}_{21}^{(1)} & \mathbf{I} \end{pmatrix}. \tag{28}$$

Effective preconditioners can be constructed from Eq. (26) with smaller number of HODLR factorization levels [89] and large low-rank factorization tolerance [90].

When l level block diagonal matrices are selected to express the impedance matrix, as in eq. (26), the expressed matrix can be inverted as an approximation of the matrix $\mathbf{Z}^{-1}$, to construct a preconditioner

$$\mathbf{M}^{-1} = \mathbf{D}_{L-l+1}^{-1}\cdots\mathbf{D}_{L-1}^{-1}\mathbf{D}_L^{-1} \tag{29}$$

The approximate matrix $\mathbf{M}^{-1}$ can be an effective preconditioning matrix for the impedance matrix. When $l = 1$, $\mathbf{M}$ is the block diagonal preconditioner, when $l = L$, it

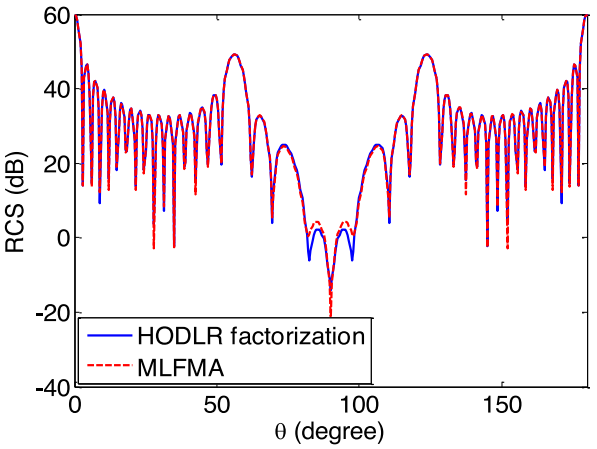

FIGURE 12. Bistatic RCS of an 16 × 16 planar array simulated with HODLR factorization and MLFMA, the average number of basis functions at leaf level is 80.

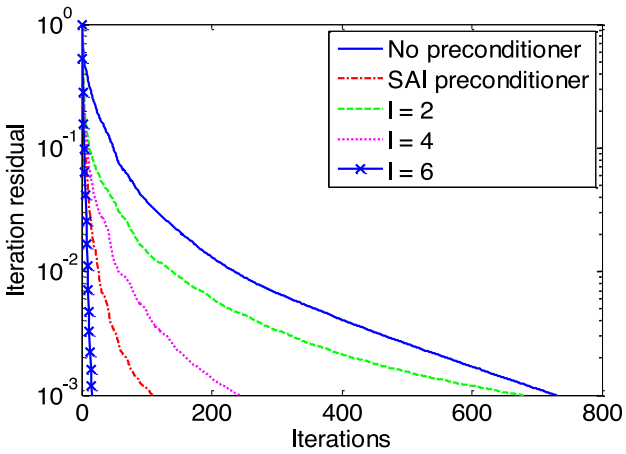

FIGURE 13. Convergence performance of the GMRES for the analysis of 16 × 16 planar array with SAI, preconditioning matrix constructed by HODLR with different levels ($l = 2, 4, 6$), and without preconditioner.

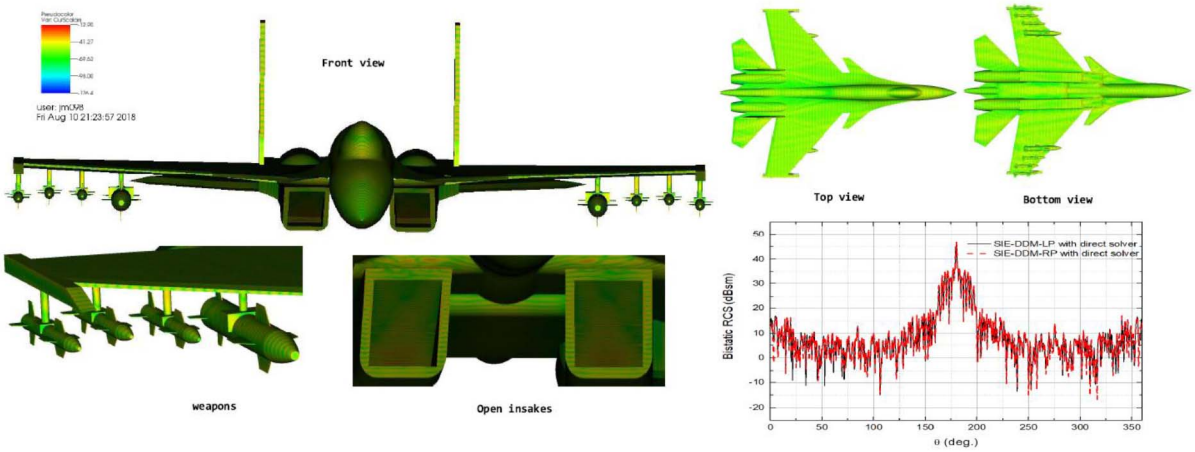

FIGURE 14. Simulated surface electric current distribution and bistatic RCS with preconditioning matrix constructed by HODLR in each subdomain. The figures are reprinted from [91].

is the full inversion matrix of $\mathbf{Z}$. When l is ranging from 1 to L, the trade-off between preconditioning performance and matrix inversion computational cost can be tested and adapted to the available resources. Both the computation time and memory complexities for the HODLR fast direct inversion are $O(Nlog^2N)$ [89], [90]. Parallel techniques based on MPI and OMP can be employed to the algorithm to further enhance the computation performance [89].

C. NUMERICAL RESULTS FOR MULTISCALE PROBLEMS

A 16×16 planar array with 333,056 unknowns is analyzed to demonstrate the performance of the proposed preconditioner. A 12-level binary tree is obtained when we fix the average number of basis functions at leaf level as 80. The simulated bistatic RCS curves with the direct HODLR factorization [89] and MLFMA [14], [15] are shown in Fig. 12, excellent agreement is observed between them. Fig. 13 shows the performance of the GMRES for the simulation of the 16×16 planar array with sparse approximate inverse (SAI) [104], [105], preconditioning matrix constructed by HODLR with different levels ($l = 2, 4, 6$), and without preconditioner. It can be clearly observed that with increasing number of levels for HODLR factorization, much better convergence can be achieved, due to the increasing ratio of the HODLR factorization parts to the whole matrix [89]. It should be noted that, with the increase of the number of levels for HODLR, the computation time and memory requirements are increasing following the computation complexity of $O(Nlog^2N)$ [87], [89]. Consequently, the proper number of levels for HODLR factorization can be chosen to achieve a balance between preconditioning performance and the matrix inversion computational cost. In our experience, l between $L/3 \sim 2L/3$ is a reasonable choice.

For realistic multiscale simulations with domain decomposition technique, the preconditioner constructed by HODLR factorization is employed in the subdomain showing poor convergence. As shown in Fig. 14, the dimensions of the aircraft are approximately 22.6 m in length, 14.2 m in wingspan, and 3.8 m in height. The incident wave is from the nose. The whole aircraft is decomposed into 13 subdomains. The number of unknowns is 3037,728 with discretization h/λ ranging from 0.03 to 0.125. For the subdomain containing open intakes cavity, with the preconditioner constructed by HODLR factorization, only 44 iterations are required to reach the residual of 0.01, showing much better performance than a block diagonal preconditioner. The surface electric current distribution and bistatic RCS results computed with left and right preconditioners are shown in Fig. 14.

VIII. CONCLUSION

In this paper, a review of the low-rank factorization method is proposed. The principles and differences between traditional low-rank based methods are illustrated. Advanced techniques to achieve better performance in terms of computation precision and efficiency are presented in detail. The source codes of NESA for the evaluation of free space Green's functions matrix produced by two coupling groups can be download freely [106].

The low-rank factorization methods are among the more promising methods in computational electromagnetics, drastical progress is achieved with the integration of mathematics and electromagnetics. They are a necessary complement to the MLFMA and FFT based methods to solve the challenges in realistic multiscale simulations.

REFERENCES

[1] R. F. Harrington, *Field Computation by Moment Methods*. New York, NY, USA: Macmillan, 1968.
[2] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propag.*, vol. 30, no. 3, pp. 409–418, May 1982.
[3] B. Z. Steinberg and Y. Leviatan, "On the use of wavelet expansions in the method of moments (EM scattering)," *IEEE Trans. Antennas Propag.*, vol. 41, no. 5, pp. 610–619, May 1993.
[4] F. X. Canning, "Solution of impedance matrix localization form of moment method problems in five iterations," *Radio Sci.*, vol. 30, no. 5, pp. 1371–1384, 1995.

[5] E. Suter and J. R. Mosig, "A subdomain multilevel approach for the efficient MoM analysis of large planar antennas," *Microw. Opt. Technol. Lett.*, vol. 26, no. 4, pp. 270–277, Aug. 2000.

[6] V. V. S. Prakash and R. Mittra, "Characteristic basis function method: A new technique for efficient solution of method of moments matrix equations," *Microw. Opt. Technol. Lett.*, vol. 36, no. 2, pp. 95–100, Jan. 2003.

[7] R. Maaskant, R. Mittra, and A. Tijhuis, "Fast analysis of large antenna arrays using the characteristic basis function method and the adaptive cross approximation algorithm," *IEEE Trans. Antennas Propag.*, vol. 56, no. 11, pp. 3440–3451, Nov. 2008.

[8] L. Matekovits, V. A. Laza, and G. Vecchi, "Analysis of large complex structures with the synthetic-functions approach," *IEEE Trans. Antennas Propag.*, vol. 55, no. 9, pp. 2509–2521, Sep. 2007.

[9] B. Zhang, G. B. Xiao, J. F. Mao, and Y. Wang, "Analyzing large-scale non-periodic arrays with synthetic basis functions," *IEEE Trans. Antennas Propag.*, vol. 58, no. 11, pp. 3576–3584, Nov. 2010.

[10] W. B. Lu, T. J. Cui, Z. G. Qian, X. X. Yin, and W. Hong, "Accurate analysis of large-scale periodic structures using an efficient sub-entire-domain basis function method," *IEEE Trans. Antennas Propag.*, vol. 52, no. 11, pp. 3078–3085, Nov. 2004.

[11] W. B. Lu, T. J. Cui, and H. Zhao, "Acceleration of fast multipole method for large-scale periodic structures with finite sizes using sub-entire-domain basis functions," *IEEE Trans. Antennas Propag.*, vol. 55, no. 2, pp. 414–421, Feb. 2007.

[12] Y. Chen and C.-F. Wang, *Characteristic Modes: Theory and Applications in Antenna Engineering*. Hoboken, NJ, USA: Wiley, 2015.

[13] L. Guan, Z. He, D. Ding, and R. S. Chen, "Efficient characteristic mode analysis for radiation problems of antenna arrays," *IEEE Trans. Antennas Propag.*, vol. 67, no. 1, pp. 199–206, Jan. 2019.

[14] L. Greengard and V. Rokhlin, "A fast algorithm for particle simulations," *J. Comput. Phys.*, vol. 73, pp. 325–348, Dec. 1987.

[15] 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, Oct. 1997.

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

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

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

[19] F. Vipiana, M. A. Francavilla, and G. Vecchi, "EFIE modeling of high-definition multiscale structures," *IEEE Trans. Antennas Propag.*, vol. 58, no. 7, pp. 2362–2374, Jul. 2010.

[20] S. Kapur and D. E. Long, "IES^3: A fast integral equation solver for efficient 3-dimensional extraction," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design*, Nov. 1997, pp. 448–455.

[21] 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, Aug. 1996.

[22] W. Hackbusch, "A sparse matrix arithmetic based on H-matrices. Part I: Introduction to H-matrices," *Computing*, vol. 62, pp. 89–108, Apr. 1999.

[23] M. Bebendorf, "Approximation of boundary element matrices," *Numerische Mathematik*, vol. 86, no. 4, pp. 565–589, Jun. 2000.

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

[25] L. Tsang, Q. Li, P. Xu, D. Chen, and V. Jandhyala, "Wave scattering with UV multilevel partitioning method: 2. Three-dimensional problem of nonpenetrable surface scattering," *Radio Sci.*, vol. 39, no. 5, pp. 1–11, Oct. 2004.

[26] M. M. Li, J. J. Ding, D. Z. Ding, Z. H. Fan, and R. S. Chen, "Multiresolution preconditioned multilevel UV Method for analysis of planar layered finite frequency selective surface," *Microw. Opt. Technol. Lett.*, vol. 52, no. 7, pp. 1530–1536, Jul. 2010.

[27] J. M. Rius, J. Parrón, A. Heldring, J. M. Tamayo, and E. Ubeda, "Fast iterative solution of integral equations with method of moments and matrix decomposition algorithm—Singular value decomposition," *IEEE Trans. Antennas Propag.*, vol. 56, no. 8, pp. 2314–2324, Aug. 2008.

[28] J. M. Tamayo, A. Heldring, and J. M. Rius, "Multilevel adaptive cross approximation (MLACA)," *IEEE Trans. Antennas Propag.*, vol. 59, no. 12, pp. 4600–4608, Dec. 2011.

[29] 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.

[30] M. Li, C. Li, C.-J. Ong, and W. Tang, "A novel multilevel matrix compression method for analysis of electromagnetic scattering from PEC targets," *IEEE Trans. Antennas Propag.*, vol. 60, no. 3, pp. 1390–1399, Mar. 2012.

[31] J.-G. Wei, Z. Peng, and J.-F. Lee, "A fast direct matrix solver for surface integral equation methods for electromagnetic wave scattering from non-penetrable targets," *Radio Sci.*, vol. 47, no. 5, pp. 1–9, Oct. 2012.

[32] M. Bebendorf and R. Venn, "Constructing nested bases approximations from the entries of non-local operators," *Numerische Mathmatik*, vol. 121, no. 4, pp. 609–635, 2012.

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

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

[35] H. Cheng *et al.*, "A wideband fast multipole method for the Helmholtz equation in three dimensions," *J. Comput. Phys.*, vol. 216, pp. 300–325, Jul. 2006.

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

[37] M. Vikram, H. Huang, B. Shanker, and T. Van, "A novel wideband FMM for fast integral equation solution of multiscale problems in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 57, no. 7, pp. 2094–2104, Jul. 2009.

[38] H. Chen, K. W. Leung, and E. K. N. Yung, "Fast directional multilevel algorithm for analyzing wave scattering," *IEEE Trans. Antennas Propag.*, vol. 59, no. 7, pp. 2546–2556, Jul. 2011.

[39] Ö. Ergül and B. Karaosmanoglu, "Broadband multilevel fast multipole algorithm based on an approximate diagonalization of the Green's function," *IEEE Trans. Antennas Propag.*, vol. 63, no. 7, pp. 3035–3041, Jul. 2015.

[40] M. Kalfa, Ö. Ergül, and V. B. Ertürk, "Error control of multiple-precision MLFMA," *IEEE Trans. Antennas Propag.*, vol. 66, no. 10, pp. 5651–5656, Oct. 2018.

[41] M. D. Altman, J. P. Bardhan, B. Tidor, and J. K. White, "FFTSVD: A fast multiscale boundary-element method solver suitable for Bio-MEMS and biomolecule simulation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 2, pp. 274–284, Feb. 2006.

[42] M. Li, R. S. Chen, H. Wang, Z. Fan, and Q. Hu, "A multilevel FFT method for the 3-D capacitance extraction," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 32, no. 2, pp. 318–322, Feb. 2013.

[43] P. Zhao, D. Q. Liu, and C. H. Chan, "A hybrid 2-D/3-D multilevel Green's function interpolation method for electrically large multilayered problems," *IEEE Trans. Antennas Propag.*, vol. 64, no. 9, pp. 3931–3942, Sep. 2016.

[44] O. M. Bucci and G. Franceschetti, "On the degrees of freedom of scattered fields," *IEEE Trans. Antennas Propag.*, vol. 37, no. 7, pp. 918–926, Jul. 1989.

[45] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Boston, MA, USA: PWS, 1996.

[46] Y. Li, H. Yang, E. R. Martin, K. L. Ho, and L. Ying, "Butterfly factorization," *Multiscale Model. Simul.*, vol. 13, no. 2, pp. 714–732, 2015.

[47] Y. Liu, X. Xing, H. Guo, E. Michielssen, P. Ghysels, and X. S. Li, "Butterfly factorization via randomized matrix-vector multiplications," 2020. [Online]. Available: arXiv:2002.03400.

[48] Y. Liu, W. Sid-Lakhdar, E. Rebrova, P. Ghysels, and X. S. Li, "A parallel hierarchical blocked adaptive cross approximation algorithm," *Int. J. High Perform. Comput. Appl.*, vol. 34, no. 4, pp. 394–408, Jul. 2020.

[49] A. Heldring, J. M. Tamayo, and J. M. Rius, "On the degrees of freedom in the interaction between sets of elementary scatterers," *presented at the 3rd Eur. Conf. Antennas Propag. (Eucap)*, Berlin, Germany, Mar. 2009, pp. 2511–2514.

[50] Z. Jiang, R.-S. Chen, Z. H. Fan, and M. Zhu, "Novel postcompression technique in the matrix decomposition algorithm for the analysis of electromagnetic problems," *Radio Sci.*, vol. 47, p. RS2003, Apr. 2012.

[51] M. Li, M. A. Francavilla, F. Vipiana, G. Vecchi, Z. Fan, and R. S. Chen, "A doubly hierarchical MoM for high-fidelity modeling of multiscale structures," *IEEE Trans. Electromagn. Compat.*, vol. 56, no. 5, pp. 1103–1111, Oct. 2014.

[52] X. Xuan, M. Yang, and Z. Jiang, "Extraction of multimode S-parameters by using a hybrid FDTD/SVD technique," *IEEE Trans. Antennas Propag.*, vol. 67, no. 3, pp. 2012–2016, Mar. 2019.

[53] X. Chen, C. Gu, J. Ding, Z. Li, and Z. Niu, "Multilevel fast adaptive cross-approximation algorithm with characteristic basis functions," *IEEE Trans. Antennas Propag.*, vol. 63, no. 9, pp. 3994–4002, Sep. 2015.

[54] X. Chen, C. Gu, A. Heldring, Z. Li, Q. Cao, "Error bound of the multilevel adaptive cross approximation (MLACA)," *IEEE Trans. Antennas Propag.*, vol. 64, no. 1, pp. 374–378, Jan. 2016.

[55] Z. Jiang, Y. Xu, Y. Sheng, and M. Zhu, "Efficient analyzing EM scattering of objects above a lossy half-space by the combined MLQR/MLSSM," *IEEE Trans. Antennas Propag.*, vol. 59, no. 12, pp. 4609–4614, Dec. 2011.

[56] C. Wang, Z. Jiang, X. Qiao, and T. Wan, "Efficient performance of MLSSM-MLFMA algorithm using adaptive grouping technique for electromagnetic problems," *IEEE Trans. Antennas Propag.*, vol. 66, no. 1, pp. 493–496, Jan. 2018.

[57] W. Hackbusch, B. Khoromskij, and S. Sauter, "On H^2–matrices," in *Lecture on Applied Mathematics*, H. Bungartz, R. Hoppe, and C. Zenger, Eds. Munich, Germany: Springer, 2000, pp. 9–29.

[58] S. Börm and W. Hackbusch, "Data-sparse approximation by adaptive H^2-matrices," *Computing*, vol. 69, pp. 1–35, Sep. 2002.

[59] S. Börm, "H^2-matrices—Multilevel methods for the approximation of integral operators," *Comput. Visual. Sci.*, vol. 7, pp. 173–181, Oct. 2004.

[60] W. Chai and D. Jiao, "An H^2-matrix-based integral-equation solver of reduced complexity and controlled accuracy for solving electrodynamic problems," *IEEE Trans. Antennas Propag.*, vol. 57, no. 10, pp. 3147–3159, Oct. 2009.

[61] E. Liberty, F. Woolfe, P.-G. Martinsson, V. Rokhlin, and M. Tygert, "Randomized algorithms for the low-rank approximation of matrices," *Proc. Nat. Acad. Sci. USA*, vol. 104, no. 51, pp. 20167–20172, 2007.

[62] X.-M. Pan, J.-G. Wei, Z. Peng, and X.-Q. Sheng, "A fast algorithm for multiscale electromagnetic problems using interpolative decomposition and multilevel fast multipole algorithm," *Radio Sci.*, vol. 47, no. 1, pp. 1–11, Feb. 2012.

[63] M. A. E. Bautista, M. A. Francavilla, P.-G. Martinsson, and F. Vipiana, "O(N) nested skeletonization scheme for the analysis of multiscale structures using the method of moments," *IEEE J. Multiscale Multiphys. Comput. Techn.*, vol. 1, pp. 139–150, Dec. 2016.

[64] X.-M. Pan and X.-Q. Sheng, "Fast solution of linear systems with many right-hand sides based on skeletonization," *IEEE Antennas Wireless Propag. Lett.*, vol. 15, pp. 301–304, 2016.

[65] D. Wu, Y.-N. Liu, Y.-M. Wu, X.-M. Pan, and X.-Q. Sheng, "Skeletonization improved calculation of electric fields by the impedance matrix of MoM," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, pp. 1108–1112, 2020.

[66] M. Li, M. A. Francavilla, F. Vipiana, G. Vecchi, and R. S. Chen, "Nested equivalent source approximation for the modeling of multiscale structures," *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3664–3678, Jul. 2014.

[67] M. Li, M. A. Francavilla, R. S. Chen, and G. Vecchi, "Wideband fast kernel-independent modeling of large multiscale structures via nested equivalent source approximation," *IEEE Trans. Antennas Propag.*, vol. 63, no. 5, pp. 2122–2134, May 2015.

[68] 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.

[69] A. Heldring, J. M. Rius, J. M. Tamayo, J. Parron, and E. Ubeda, "Fast direct solution of method of moments linear system," *IEEE Trans. Antennas Propag.*, vol. 55, no. 11, pp. 3220–3228, Nov. 2007.

[70] 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.

[71] A. Heldring, J. M. Tamayo, E. Ubeda, and J. M. Rius, "Accelerated direct solution of the method-of-moments linear system," *Proc. IEEE*, vol. 101, no. 2, pp. 364–371, Feb. 2013.

[72] Z. Jiang, Y. Sheng, and S. Shen, "Multilevel fast multipole algorithm-based direct solution for analysis of electromagnetic problems," *IEEE Trans. Antennas Propag.*, vol. 59, no. 9, pp. 3491–3494, Sep. 2011.

[73] T. Wan, Z. N. Jiang, and Y. J. Sheng, "Hierarchical matrix techniques based on matrix decomposition algorithm for the fast analysis of planar layered structures," *IEEE Trans. Antennas Propag.*, vol. 59, no. 11, pp. 4132–4141, Nov. 2011.

[74] W. Chai and D. Jiao, "An LU decomposition based direct integral equation solver of linear complexity and higher-order accuracy for large-scale interconnect extraction," *IEEE Trans. Adv. Packag.*, vol. 33, no. 4, pp. 794–803, Nov. 2010.

[75] H. Shao, J. Hu, H. Guo, F. Ye, W. Lu, and Z. Nie, "Fast simulation of array structures using T-EPA with hierarchical LU decomposition," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 1556–1559, 2012.

[76] T. Wan, Q. I. Dai, and W. C. Chew, "Fast low-frequency surface integral equation solver based on hierarchical matrix algorithm," *Progr. Electromagn. Res.*, vol. 161, pp. 19–33, Mar. 2018.

[77] T. Wan, Q. Zhang, T. Hong, D. Z. Ding, Z. H. Fan, and R. S. Chen, "Fast analysis of three-dimensional electromagnetic problems using dual-primal finite-element tearing and interconnecting method combined with H-matrix technique," *IET Microw. Antennas Propag.*, vol. 59, no. 11, pp. 4132–4141, Jun. 2015.

[78] T. Wan, B. Tang, and M. Li, "An iteration-free domain decomposition method for the fast finite element analysis of electromagnetic problems," *IEEE Trans. Antennas Propag.*, vol. 68, no. 1, pp. 400–410, Jan. 2020.

[79] 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 Techn.*, vol. 59, no. 10, pp. 2404–2421, Oct. 2011.

[80] W. Chai and D. Jiao, "Direct matrix solution of linear complexity for surface integral-equation-based impedance extraction of complicated 3-D structures," *Proc. IEEE*, vol. 101, no. 2, pp. 372–388, Feb. 2013.

[81] S. Omar and D. Jiao, "A linear complexity direct volume integral equation solver for full-wave 3-D circuit extraction in inhomogeneous materials," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 3, pp. 897–912, Mar. 2015.

[82] 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 Trans. Microw. Theory Techn.*, vol. 66, no. 1, pp. 35–48, Jan. 2018.

[83] M. Ma and D. Jiao, "Direct solution of general H^2 -matrices with controlled accuracy and concurrent change of cluster bases for electromagnetic analysis," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 6, pp. 2114–2127, Jun. 2019.

[84] S. Chandrasekaran *et al.*, "Some fast algorithms for sequentially semiseparable representations," *SIAM J. Matrix Anal. Appl.*, vol. 27, no. 2, pp. 341–364, 2006.

[85] J. Xia, S. Chandrasekaran, M. Gu, and X. S. Li, "Fast algorithms for hierarchically semiseparable matrices," *Number. Linear Algebra Appl.*, vol. 17, no. 6, pp. 953–976, Dec. 2010.

[86] M. Ma and D. Jiao, "New HSS-factorization and inversion algorithms with exact arithmetic for efficient direct solution of large-scale volume integral equations," in *Proc. IEEE Int. Symp. Antennas Propag. (APSURSI)*, Fajardo, PR, USA, Jul. 2016, pp. 1567–1568.

[87] S. Ambikasaran and E. Darve, "An O($N\log N$) fast direct solver for partial hierarchically semi-separable matrices," *J. Sci. Comput.*, vol. 57, no. 3, pp. 477–501, 2013.

[88] X. Chen, C. Gu, Z. Li, and Z. Niu, "Accelerated direct solution of electromagnetic scattering via characteristic basis function method with Sherman-Morrison-Woodbury formula-based algorithm," *IEEE Trans. Antennas Propag.*, vol. 64, no. 10, pp. 4482–4486, Oct. 2016.
[89] K. Wang, M. Li, D. Ding, and R. S. Chen, "A parallelizable direct solution of integral equation methods for electromagnetic analysis," *Eng. Anal. Boundary Elements*, vol. 85, pp. 158–164, Dec. 2017.
[90] Z. Rong *et al.*, "Fast direct solution of integral equations with modified HODLR structure for analyzing electromagnetic scattering problems," *IEEE Trans. Antennas Propag.*, vol. 67, no. 5, pp. 3288–3296, May 2019.
[91] M. Jiang *et al.*, "SIE-DDM based on a hybrid direct–iterative approach for analysis of multiscale problems," *IEEE Trans. Antennas Propag.*, vol. 67, no. 12, pp. 7440–7451, Dec. 2019.
[92] J. Laviada, R. Mittra, M. R. Pino, and F. Las-Heras, "On the convergence of the ACA," *Microw. Opt. Technol. Lett.*, vol. 51, no. 10, pp. 2458–2460, Oct. 2009.
[93] A. Heldring, E. Ubeda, and J. M. Rius, "On the convergence of the ACA algorithm for radiation and scattering problems," *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3806–3809, Jul. 2014.
[94] A. Heldring, E. Ubeda, and J. M. Rius, "Improving the accuracy of the adaptive cross approximation with a convergence criterion based on random sampling," *IEEE Trans. Antennas Propag.*, vol. 69, no. 1, pp. 347–355, Jan. 2021, doi: 10.1109/TAP.2020.3010857.
[95] 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 Propag. Mag.*, vol. 35, no. 1, pp. 84–89, Feb. 1993.
[96] N. Halko, P. G. Martinsson, and J. A. Tropp, "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM Rev.*, vol. 53, no. 2, pp. 217–288, 2011.
[97] P. G. Martinsson, "A fast randomized algorithm for computing a hierarchically semiseparable representation of a matrix," *SIAM J. Matrix Anal. Appl.*, vol. 32, no. 4, pp. 1251–1274, 2011.
[98] F. P. Andriulli, F. Vipiana, and G. Vecchi, "Hierarchical bases for non-hierarchic 3-D triangular meshes," *IEEE Trans. Antennas Propag.*, vol. 56, no. 8, pp. 2288–2297, Aug. 2008.
[99] M. A. Francavilla, F. Vipiana, G. Vecchi, and D. R. Wilton, "Hierarchical fast MoM solver for the modeling of large multiscale wire-surface structures," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 1378–1381, 2012.
[100] M. Li, M. A. Francavilla, R. Chen, and G. Vecchi, "Nested equivalent source approximation for the modeling of penetrable bodies," *IEEE Trans. Antennas Propag.*, vol. 65, no. 2, pp. 954–959, Feb. 2017.
[101] M. Li, M. A. Francavilla, D. Ding, R. Chen, and G. Vecchi, "Mixed-form nested approximation for wideband multiscale simulations," *IEEE Trans. Antennas Propag.*, vol. 66, no. 11, pp. 6128–6136, Nov. 2018.
[102] B. Engquist and L. Ying, "Fast directional multilevel algorithms for oscillatory kernels," *SIAM J. Sci. Comput.*, vol. 29, no. 4, pp. 1710–1737, 2007.
[103] M. Bebendorf, C. Kuske, and R. Venn, "Wideband nested cross approximation for Helmholtz problems," *Numerische Mathematik*, vol. 130, no. 1, pp. 1–34, 2015.
[104] M. Benzi and M. Tuma, "A sparse approximate inverse preconditioner for nonsymmetric linear systems," *J. Sci. Comput.*, vol. 19, no. 3, pp. 968–994, 1998.
[105] P. L. Rui and R. S. Chen, "An efficient sparse approximate inverse preconditioning for FMM implementation," *Microw. Opt. Technol. Lett.*, vol. 49, no. 7, pp. 1746–1750, Jul. 2007.
[106] *mengmeng-nanjing/Nested-Equivalence-Source-Approximation*. Accessed: Feb. 2021. [Online]. Available: https://github.com/mengmeng-nanjing/Nested-Equivalence-Source-Approximation

MENGMENG LI (Senior Member, IEEE) received the B.S. degree (Hons.) in physics from Huaiyin Normal College, Huai'an, China, in 2007, and the Ph.D. degree in electromagnetic field and microwave technology from the Nanjing University of Science and Technology, Nanjing, China, in 2014.

From 2012 to 2014, he was a visiting student with the Electronics Department, Politecnico di Torino, Turin, Italy, and also with the Antenna and EMC Laboratory (LACE), Istituto Superiore Mario Boella, Turin, where he carried out fast solver for multiscale simulations. Since 2014, he has been with the Department of Communication Engineering, Nanjing University of Science and Technology, where he has been an Assistant Professor, an Associate Professor, and a Professor since 2020. In 2017, he was a Visiting Scholar with Pennsylvania State University, Pennsylvania, PA, USA. His current research interests include fast solver algorithms, computational electromagnetic solvers for circuits, signal integrity analysis, and multiscale simulations. He was a recipient of the Doctoral Dissertation Award of Jiangsu Province in 2016, the Young Scientist Award at the ACES-China Conference in 2019, and five student paper/contest awards at the international conferences with the students. He is an active reviewer for many IEEE journals and conferences. He is an Associate Editor of the *IEEE Antennas and Propagation Magazine*, IEEE OPEN JOURNAL OF ANTENNAS AND PROPAGATION, and IEEE ACCESS, and a Guest Editor of IEEE OPEN JOURNAL OF ANTENNAS AND PROPAGATION.

DAZHI DING received the B.Sc. and Ph.D. degrees in electromagnetic field and microwave technique from the Nanjing University of Science and Technology (NJUST), Nanjing, China, in 2002 and 2007, respectively.

In 2005, he was with the Center of wireless Communication, City University of Hong Kong, Hong Kong, as a Research Assistant. He joined the Department of Electrical Engineering, NJUST, where he became a Lecturer in 2007. In 2014, he was promoted to Full Professor with NJUST, where he was appointed as the Head of the Department of Communication Engineering, in September 2014. He has authored or coauthored over 30 technical articles. He has authored or coauthored more than 80 articles. His current research interests include computational electromagnetics and electromagnetic scattering and radiation. He was a recipient of the National Excellent Youth Fund by the National Science Foundation of China in 2020.

ALEXANDER HELDRING was born in Amsterdam, The Netherlands, in 1966. He received the M.S. degree in applied physics and the Ph.D. degree in electrical engineering from the Delft University of Technology, Delft, The Netherlands, in 1993 and 2002, respectively.

He is currently working as an Associate Professor with the Telecommunications Department, Universitat Politècnica de Catalunya, Barcelona, Spain. He has authored over 30 articles in international journals and 100 in international conference proceedings. His special research interest includes fast integral equation methods for electromagnetic problems.

JUN HU (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electromagnetic field and microwave technique from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1995, 1998, and 2000, respectively.

In 2001, he was a Research Assistant with the Center of Wireless Communication, City University of Hong Kong, Hong Kong. In 2010, he was a Visiting Scholar with the Electro Science Laboratory, Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH, USA. He was a Visiting Professor with the City University of Hong Kong, Hong Kong, in 2011. Since 2017, he has been a Vice President of UESTC, where he is currently a Full Professor with the School of Electronic Science and Engineering. He has authored or coauthored over 300 technical papers. His current research interests include integral equation methods in computational electromagnetics and electromagnetic scattering and radiation.

Dr. Hu was a recipient of the 2004 Best Young Scholar Paper Prize of Chinese Radio Propagation Society, the 2014 National Excellent Youth Fund by the NSFC, awarded as Chang Jing Scholar in 2016, and many best student papers awards. He was a co-recipient of the 2018 IEEE AP-S Sergei A. Schelkunoff Transaction Paper Award. He served as a Chairman for Student Activities Committee of the IEEE Cheng du Section from 2010 to 2016, and the IEEE Chengdu AP/EMC Joint Chapter from 2014 to 2016.

RUSHAN CHEN (Senior Member, IEEE) was born in Taizhou, China. He received the B.Sc. and M.Sc. degrees from the Department of Radio Engineering, Southeast University, Nanjing, China, in 1987 and 1990, respectively, and the Ph.D. degree from the Department of Electronic Engineering, City University of Hong Kong, Hong Kong, in 2001.

He joined the Department of Electrical Engineering, Nanjing University of Science and Technology (NJUST), Nanjing, where he became a Teaching Assistant in 1990 and a Lecturer in 1992. Since 1996, he has been a Visiting Scholar with the Department of Electronic Engineering, City University of Hong Kong, first as a Research Associate, and then a Senior Research Associate in 1997, a Research Fellow in 1998, and a Senior Research Fellow in 1999. In 1999, he was also a Visiting Scholar with Montreal University, Montreal, QC, Canada. In 1999, he was promoted to a Full Professor and the Associate Director of the Microwave and Communication Research Center, NJUST, where he was the Head of the Department of Communication Engineering in 2007. In 2008, he was a Chang-Jiang Professor under the Cheung Kong Scholar Program awarded by the Ministry of Education, China. In 2009, he was the Dean of the School of Communication and Information Engineering, Nanjing Post and Communications University, Nanjing. In 2011, he was the Vice-Dean of the School of Electrical Engineering and Optical Technique, NJUST. He has authored or coauthored more than 200 articles, including over 140 articles in international journals. His current research interests include microwave/millimeter-wave systems, measurements, antenna, RF-integrated circuits, and computational electromagnetics. He was a recipient of the 1992 Third-Class Science and Technology Advance Prize given by the National Military Industry Department of China, the 1993 Third-Class Science and Technology Advance Prize given by the National Education Committee of China, the 1996 Second-Class Science and Technology Advance Prize given by the National Education Committee of China, the 1999 First-Class Science and Technology Advance Prize given by Jiangsu Province, the 2001 Second-Class Science and Technology Advance Prize, and the Foundation for China Distinguished Young Investigators presented by the National Science Foundation of China in 2003. He serves as a Reviewer for many technical journals, such as the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNOLOGY, and *Chinese Physics*. He is currently an Associate Editor of the *International Journal of Electronics*. He is a Senior Member of the Chinese Institute of Electronics and the Vice-President of the Microwave Society of CIE and the IEEE MTT/APS/EMC Nanjing Chapter. He was a member of the Electronic Science and Technology Group by Academic Degree Commission of the State Council in 2009.

GIUSEPPE VECCHI (Fellow, IEEE) received the Laurea and Ph.D. (Dottorato di Ricerca) degrees in electronic engineering from the Politecnico di Torino, Torino, Italy, in 1985 and 1989, respectively, with doctoral research carried out partly with Polytechnic University, Farmingdale, NY, USA.

He was a Visiting Scientist with Polytechnic University from 1989 to 1990, and since 1990 he is with the Department of Electronics, Politecnico di Torino, where he has been Assistant Professor from 1992 to 2000, and a Professor. He was a Visiting Scientist with the University of Helsinki, Helsinki, Finland, in 1992, and has been an Adjunct Faculty with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, IL, USA, from 1997 to 2011. Since 2015 he serves as the Director of the Antenna and EMC Lab (LACE), Politecnico di Torino, Torino, Italy. His current research activities concern analytical and numerical techniques for design, measurement and diagnostics of antennas and devices, medical applications, and imaging.

Prof. Vecchi is a Fellow of the IEEE, and Member of the Board of the European School of Antennas. He has been an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, Chairman of the IEEE Antennas & Propagation Society, IEEE Microwave Theory and Techniques Society, IEEE Transactions on Electron Devices, and Italian joint Chapter, and Member of the IEEE-APS Educational Committee.