An Efficient and Error-Controllable Angular Sweep Algorithm for Electromagnetic Wave Scattering and Its Analysis

Chung Hyun Lee, *Student Member, IEEE*, Joseph D. Kotulski, *Life Senior Member, IEEE*, Vinh Q. Dang, *Member, IEEE*, and Jin-Fa Lee, *Fellow, IEEE*

***Abstract*— The angular sweep of electromagnetic wave scattering is formulated as a matrix equation with multiple right-hand sides (RHSs). Although the low-rank approximation of an RHS matrix is a popular choice for reducing the computational costs of multiple RHSs, only a small amount of research has been conducted to explore how this approximation impacts the solution quality. Furthermore, there has not been sufficient research on the quality of the solution as a function of the accuracy of the iterative solver. We present an error analysis of the approximated solution considering both the reduced number of RHSs and the tolerance of the iterative solver. Based on the error analysis, a new angular sweep algorithm is proposed with fine-tuned tolerances of the iterative solver for individual singular vectors. The different tolerances for each singular vector increase the efficiency of the proposed algorithm. Another benefit of the proposed algorithm is that the error can be bounded by a user-defined global tolerance. In addition, a variant of the generalized conjugate residual method for multiple RHSs is introduced to accelerate iterative solvers. Finally, numerical validation is conducted with three examples in which the discontinuous Galerkin surface integral equation method is applied. The experiments support two conclusions: tight upper and lower bounds of the solution error exist, and fine-tuning the tolerances reduces the computational costs.**

***Index Terms*— Angular response, discontinuous Galerkin surface integral (SIE) equation, electromagnetic wave scattering, error analysis, Krylov space method, multiple right-hand sides (RHSs).**

I. Introduction

FULL-WAVE electromagnetic simulation has been used in a variety of applications [1], [2], [3] due to the development of successful algorithms and the expansion of processing capability over the past few decades. For these simulation methods, multiple angular responses of electromagnetic wave scattering are a classic challenge in the area of computational electromagnetics. A right-hand side (RHS) matrix can be used to construct multiple incident fields, with each column vector representing an excitation vector for a particular incident angle. Standard direct solvers, such as lower–upper (LU) decomposition [4], are favored for solving multiple RHSs; in general, however, because of their complexity, advanced methods are necessary. The hierarchical matrices approach [5] is one of the direct solvers for addressing the abovementioned problems, even for electrically large targets. Zhou and Jiao [6] and Guo et al. [7] applied hierarchical matrix methods in combination with the finite-element method (FEM) and integral equation, respectively, although the computational complexity of the work in the latter article is expected to be $\mathcal{O}(N^{1.5}\log N)$. Alternatively, block Krylov subspace methods are a viable option because several RHSs can be solved simultaneously. Various examples are introduced, such as the block generalized minimal residual (BGMRES) [8] and the block generalized conjugate residual with optimal truncation [BGCROT(m,k)] [9]. Another effort to speed up iterative solvers for multiple RHSs is the generalized conjugate residual with inner orthogonality and deflation restarting [GCRO-DR(m,k)] [10], [11].

In addition to block iterative solvers, other efforts are underway, which reduce the dimension of multiple RHSs itself. Low-rank approximation with singular value decomposition (SVD) [12], asymptotic waveform evaluation (AWE) [13], [14], and model-based parameter estimation (MBPE) [15], [16] are well-known methods. In addition, Peng et al. [11] introduced a novel method that uses Fourier harmonics to construct excitation matrices. This method has two advantages: the accuracy of the excitation matrix can be estimated and the sampling size can be adaptively increased. However, this approach is limited because the tolerance of SVD is not systematically chosen, and also, the SVD is computationally expensive. Adaptive cross approximation (ACA) is one of the remedies in [17] and [18]. In particular, Kazempour and Gürel [17] applied a recompressed ACA (RACA) to further compress the suboptimal factorization of ACA, although this can be criticized because of the necessity of the a priori choice of both tolerances of ACA and iterative solver. In [19], interpolative decomposition (ID) was used to shrink the dimension of incident vectors with a given tolerance. It is important for this article to discuss the error control of the algorithm considering both the accuracy of ID and indirect analysis of the

Manuscript received 9 March 2022; revised 31 October 2022; accepted 13 December 2022. Date of publication 9 January 2023; date of current version 6 March 2023. This work was supported by Sandia National Laboratories (a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc.) through the U.S. Department of Energy's National Nuclear Security Administration under Contract DE-NA-0003525. *(Corresponding author: Chung Hyun Lee.)*

Chung Hyun Lee and Jin-Fa Lee are with the Department of Electrical Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: lee.4542@osu.edu; lee.1863@osu.edu).

Joseph D. Kotulski and Vinh Q. Dang are with Sandia National Laboratories, Albuquerque, NM 87185 USA (e-mail: jdkotul@sandia.gov; vqdang@sandia.gov).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TAP.2022.3233426.

Digital Object Identifier 10.1109/TAP.2022.3233426

relationship between surface currents and radar cross sections. Nonetheless, the error analysis was not rigorous because the numerical error of the iterative solver for each skeleton vector was overlooked. Recently, experimental approaches [20], [21] have been proposed with compressive sensing [22]. These articles show that compressive sensing methods can reduce the dimensionality of multiple incident field vectors, although the practical advantages of these methods are debatable due to a lack of both error analysis and systematic selection of the initial number of incident angles.

Based on the proper selection of angular samples and the low-rank approximation of the RHS matrix, this article introduces an optimal and error-controllable algorithm for fast angular sweeps. To minimize the computational costs, distinct tolerances are chosen automatically by the algorithm for the reduced RHS vectors rather than using the same tolerance for all vectors. One major benefit of the algorithm is that the error of the recovered solution is bounded by the user-defined tolerance, as demonstrated later. Also, the proposed algorithm can be applied to both dense and sparse matrices, although surface integral (SIE) methods are used as an example in this article. In addition, a simple variant of the block generalized conjugate residual (GCR) is also introduced to speed up the iterative solution procedure.

There are three major contributions of this article:

1) the various tolerances for the iterative solver;
2) the theoretical error bound due to the compression;
3) the precise error controllability according to the error analysis.

The following is a breakdown of how this article is structured. In Section II, the preliminaries are presented, followed by extensive explanations of the algorithm. Also, the main results are presented in Section II-D (Theorem 1) with error analysis. In Section III, we present the numerical validation of error bounds as well as the efficiency of our method with three different targets, one of which is electrically large. Finally, the conclusions are presented in Section IV.

II. Main Idea

A. Preliminaries

In this article, a boldface capital letter and a boldface small letter denote a matrix and a vector, respectively. $\mathcal{I}_j$ is an index set defined by $\mathcal{I}_j := \{1, \ldots, j\}$. In addition, a matrix norm can be defined by the vector norm [23], [24], i.e., for an arbitrary matrix $\mathbf{W} \in \mathbb{C}^{N\times M}$

$$\|\mathbf{W}\| = \sup_{\mathbf{x}\neq\mathbf{0}, \mathbf{x}\in\mathbb{C}^M} \frac{\|\mathbf{W}\mathbf{x}\|}{\|\mathbf{x}\|}. \tag{1}$$

Unless otherwise specified, the matrix norm refers to this vector-induced matrix norm.

B. Problem Statements

To begin, consider the following definition.

Definition 1: A (preconditioned) matrix equation with M normalized RHSs ($\|\mathbf{b}_i\| = 1$, $\forall i \in \mathcal{I}_M$)

$$\mathbf{A}\mathbf{X} = \mathbf{B} \tag{2}$$

is said to be solved with a given tolerance δ if

$$\epsilon_r := \frac{\|\mathbf{E}\|}{\|\mathbf{B}\|} = \frac{\|\mathbf{B} - \mathbf{A}\overline{\mathbf{X}}\|}{\|\mathbf{B}\|} \leq \delta. \tag{3}$$

$\overline{\mathbf{X}}$ is a numerical solution of $\mathbf{X}$.

In Definition 1, $\mathbf{A} \in \mathbb{C}^{N\times N}$ is a system matrix expanded with N basis functions. An excitation matrix $\mathbf{B}$ is defined by $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_M] \in \mathbb{C}^{N\times M}$, and each $\mathbf{b}_i$ is an N-dimensional column vector of the plane wave excitation at a particular incident angle. Both $\mathbf{A}$ and $\mathbf{B}$ can be regarded as matrices after applying preconditioners. In addition, $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_M] \in \mathbb{C}^{N\times M}$ is an unknown solution matrix whose column vectors, $\mathbf{x}_i$, are an N-dimensional unknown surface current vector for the corresponding incident vector, $\mathbf{b}_i$. ϵ_r is a relative error used to evaluate the accuracy of the solution, and δ is a predetermined global tolerance. The iterative solver uses (3) as the convergence criteria.

The objective of our proposed algorithm is to find a numerical solution $\overline{\mathbf{X}} = [\overline{\mathbf{x}}_1, \ldots, \overline{\mathbf{x}}_M] \in \mathbb{C}^{N\times M}$ such that (3) is satisfied. Ideally, the best choice for estimating the error of the solution ($\mathbf{E}$) is that every iteration computes the vector-induced matrix norm; nevertheless, the computational cost is excessive because it requires SVD. As a result, we later propose a method for estimating the matrix norm with respect to the tolerances of the iterative solver for each singular vector.

C. Proposed Algorithm

The algorithm is made up of the following four steps.

1) Computation of Incident Vectors: For a given target, the number of samples shall be appropriately determined to ensure to capture the desired angular responses and to minimize the computation. According to various articles [19], [25], [26], the angular steps can be limited by

$$\Delta\varphi, \Delta\theta \leq \frac{\pi}{k_0 D + 1.8(d_0)^{\frac{2}{3}}(k_0 D)^{\frac{1}{3}}} \tag{4}$$

where d_0 is the number of digits of the accuracy, k_0 is the free-space wavenumber, and D is the length of a cube that encloses the target. Also, the unit of both $\Delta\varphi$ and $\Delta\theta$ is radians. Then, we can determine the number of incident vectors by

$$M = \left(\left\lceil \frac{|\varphi_{end} - \varphi_{start}|}{\Delta\varphi} \right\rceil + 1\right)\left(\left\lceil \frac{|\theta_{end} - \theta_{start}|}{\Delta\theta} \right\rceil + 1\right) \tag{5}$$

where $\lceil\cdot\rceil$ is the ceiling function and φ_{start} and φ_{end} are the start and stop azimuthal angles, respectively. Similarly, θ_{start} and θ_{end} are denoted as the start and stop elevation angles, respectively. We assume that all incident vectors shall be normalized by its vector norm; therefore, $\|\mathbf{b}_i\| = 1, \forall i \in \mathcal{I}_M$.

2) Low-Rank Approximation: Even with the optimal sampling number, the incident vectors can be easily rank-deficient, as shown in Section III-A. Therefore, the second step is the compression of the given RHS matrix to check the linear dependencies. The proposed algorithm can be used with various low-rank approximations. One common approach is the SVD to produce a low-rank approximation of the excitation matrix. Namely, $\mathbf{B}$ is represented by

$$\mathbf{B} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H. \tag{6}$$

Here, $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_M] \in \mathbb{C}^{N\times M}$ is a matrix with left singular vectors and $\mathbf{\Sigma}$ is a diagonal matrix with descending order singular values, $\mathbf{\Sigma} = diag\{\sigma_1, \ldots, \sigma_M\} \in \mathbb{R}^{M\times M}$. In addition, $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_M] \in \mathbb{C}^{M\times M}$ is a matrix with right singular vectors. The superscript "H" denotes the conjugate transpose. When determining appropriate k values, one can compress the original excitation matrix into a rank k matrix having k singular vectors with a preferred accuracy. In the proposed algorithm, a relative singular value is used to choose k. Namely, a numerical rank k [27] is identified such that

$$k = \min\left\{r : \frac{\sigma_{r+1}}{\sigma_1} < \delta\right\}. \tag{7}$$

As shown later in Theorem 1, the relative singular value, σ_{k+1}/σ_1, is a lower bound of ϵ_r in (3).

To minimize the computational cost, a randomized algorithm, e.g., principal component analysis (PCA) [28], [29], [30], can be utilized. Although the conventional SVD subroutine is used to validate the theoretical error bounds in Section III-C, the PCA process is applied in Section III-E using the subsampled randomized Fourier transform (SRFT) presented in [30] due to its favorable complexity, with $\mathcal{O}(NM \log L)$. As mentioned before, the ACA-related approach [17], [18] is an alternative choice.

3) Solve With Singular Vectors: The next step is to solve for $\mathbf{Z}_{1:k} = [z_1, \ldots, z_k] \in \mathbb{C}^{N\times k}$ with k singular vectors and zero-out all other solution vectors. Namely, z_i with $\mathbf{u}_i$ is solved as follows:

$$\mathbf{A}z_i = \mathbf{u}_i, \ \forall i \in \mathcal{I}_k \tag{8}$$

and set

$$z_i = \mathbf{0}, \ \forall i \in \mathcal{I}_M - \mathcal{I}_k. \tag{9}$$

There is nothing to do with (9) when implementing the algorithm because solving only (8) implies that (9) has already been satisfied. A critical parameter for solving (8) is the stopping criteria if an iterative solver is chosen. This tolerance is essential because it is directly related to the desired solution quality without paying excessive computational cost. To solve (8), we set the tolerance of each singular vector as

$$\delta_i := \frac{1}{\sigma_i\sqrt{k}}(\sigma_1\delta - \sigma_{k+1}) \tag{10}$$

for $i \in \mathcal{I}_k$. The most important idea for the proposed algorithm is that each singular vector is solved with "different tolerance" δ_i in (10). The convergence criteria can be relaxed, instead of having the same tolerance for all singular vectors [11], by assigning the larger tolerance to the vectors with smaller singular values. In other words, the computational costs can be reduced compared to solving the same tolerance for all singular vectors. The advantage of choosing different tolerances is shown in Section III-D.

4) Recover Original Solution: Finally, the solution can be recovered with

$$\mathbf{X} \approx \overline{\mathbf{X}} = \overline{\mathbf{Z}}_{1:k}\mathbf{\Sigma}_{1:k}\mathbf{V}_{1:k}^H \tag{11}$$

where $\overline{\mathbf{Z}}_{1:k} = [\overline{z}_1, \ldots, \overline{z}_k] \in \mathbb{C}^{N\times k}$ is a numerical solution matrix of $\mathbf{Z}_{1:k}$, $\mathbf{\Sigma}_{1:k}$ is a $k \times k$ diagonal matrix with the first k singular values of $\mathbf{\Sigma}$, and $\mathbf{V}_{1:k}$ is an $M \times k$ matrix with the first k right singular vectors of $\boldsymbol{V}$. The numerical solution (11) guarantees (3) according to the error bound derivation in Section II-D.

The proposed algorithm is summarized in Algorithm 1.

Algorithm 1 Efficient and Error-Controllable Algorithm

Given: $\mathbf{A}^{N\times N}$ and δ
Result: $\overline{\mathbf{X}}$

Step 1: Determine M with (5)
Step 2: $\mathbf{B} \leftarrow$ Normalized incident vectors
Step 3: $[\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}^H] \leftarrow$ Apply SVD or PCA to $\mathbf{B}$
Step 4: Determine k with (7)
Step 5: Solve with k singular vectors
  for $i \leftarrow 1$ **to** k **do**
    Solve $\mathbf{A}z_i = \mathbf{u}_i$ for z_i with a tolerance δ_i in (10)
  end for
Step 6: $\overline{\mathbf{X}} \leftarrow \overline{\mathbf{Z}}_{1:k}\mathbf{\Sigma}_{1:k}\mathbf{V}_{1:k}^H$

D. Error Bounds

In this section, the analysis of the error bounds of ϵ_r is presented. With Lemma 1, the error bounds are stated and proved in Theorem 1 that the maximum relative error is between the predetermined tolerance δ and the truncation criteria in (7). In addition, Corollary 1 provides an insight into the numerical error for individual incident vectors with the given tolerance. The derivation starts from the following two notations for $i \in \mathcal{I}_M$:

$$\mathbf{B}_i := \mathbf{u}_i\sigma_i\mathbf{v}_i^H \tag{12}$$

and

$$\mathbf{X}_i := \overline{z}_i\sigma_i\mathbf{v}_i^H. \tag{13}$$

Note that $\mathbf{B}_i \in \mathbb{C}^{N\times M}$ and $\mathbf{X}_i \in \mathbb{C}^{N\times M}$ are rank-one matrices. Also,

$$\sum_{i\in\mathcal{I}_M}\mathbf{B}_i = \mathbf{B} \text{ and } \sum_{i\in\mathcal{I}_M}\mathbf{X}_i = \overline{\mathbf{X}}. \tag{14}$$

Lemma 1: Let (8) be solved with δ_i defined in (10). Then,

$$\left\|\sum_{i\in\mathcal{I}_k}(\mathbf{B}_i - \mathbf{A}\mathbf{X}_i)\right\| \le \sigma_1\delta - \sigma_{k+1}. \tag{15}$$

Proof: For any index $i \in \mathcal{I}_k$, let

$$\boldsymbol{\xi}_i := \mathbf{u}_i - \mathbf{A}\overline{z}_i. \tag{16}$$

Consider solving the matrix equation $\mathbf{A}z_i = \mathbf{u}_i$ with the tolerance δ_i defined in (10). When any iterative solver converges to target tolerance δ_i, it is trivial that

$$\|\boldsymbol{\xi}_i\| = \|\mathbf{u}_i - \mathbf{A}\overline{z}_i\| \le \delta_i. \tag{17}$$

With (16), one can write

$$\sum_{i\in\mathcal{I}_k}(\mathbf{B}_i - \mathbf{A}\mathbf{X}_i) = \begin{bmatrix}\boldsymbol{\xi}_1 & \cdots & \boldsymbol{\xi}_k\end{bmatrix}\mathbf{\Sigma}_{1:k}\mathbf{V}_{1:k}^H \tag{18}$$

where $\mathbf{\Sigma}_{1:k}$ and $\mathbf{V}_{1:k}$ are defined in (11). Then, the norm of (18) is bounded by

$$\begin{aligned}\left\|\sum_{i\in\mathcal{I}_k}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\| &\le \left\|\left[\frac{\boldsymbol{\xi}_1}{\|\boldsymbol{\xi}_1\|}\ \cdots\ \frac{\boldsymbol{\xi}_k}{\|\boldsymbol{\xi}_k\|}\right]\right\| \\ &\quad\times\left\|\operatorname{diag}\left(\sigma_1\|\boldsymbol{\xi}_1\|,\cdots,\sigma_k\|\boldsymbol{\xi}_k\|\right)\right\|\left\|\mathbf{V}_{1:k}^H\right\| \\ &\le \left\|\left[\frac{\boldsymbol{\xi}_1}{\|\boldsymbol{\xi}_1\|}\ \cdots\ \frac{\boldsymbol{\xi}_k}{\|\boldsymbol{\xi}_k\|}\right]\right\|_F \max_{i\in\mathcal{I}_k}\left(\sigma_i\|\boldsymbol{\xi}_i\|\right)\left\|\mathbf{V}_{1:k}^H\right\| \\ &= \sqrt{k}\max_{i\in\mathcal{I}_k}\left(\sigma_i\|\boldsymbol{\xi}_i\|\right) \\ &\le \sqrt{k}\max_{i\in\mathcal{I}_k}(\sigma_i\delta_i)=\sigma_1\delta-\sigma_{k+1} \end{aligned} \tag{19}$$

where $\|\cdot\|_F$ is the Frobenius norm. The second inequality is valid due to the fact that $\|\mathbf{W}\|\le\|\mathbf{W}\|_F$, for an arbitrary matrix $\mathbf{W}\in\mathbb{C}^{N\times M}$ [23], [24]. Finally, the last equality of (19) can be proven by replacing δ_i in (10). □

Theorem 1: Let (8) be solved with δ_i defined in (10) and set (9). Then,

$$\frac{\sigma_{k+1}}{\sigma_1}\le\frac{\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\|}{\|\mathbf{B}\|}\le\delta. \tag{20}$$

Proof: Multiplying a normalization factor $\|\mathbf{B}\|=\sigma_1$ to both sides, (20) can be rewritten as

$$\sigma_{k+1}\le\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\|\le\sigma_1\delta. \tag{21}$$

For the upper bound, the following equation provides a starting point:

$$\begin{aligned}\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\| &= \left\|\sum_{i\in\mathcal{I}_M}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\| \\ &\le \underbrace{\left\|\sum_{i\in\mathcal{I}_k}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\|}_{\text{Controllable with }\delta_i}+\underbrace{\left\|\sum_{i\in\mathcal{I}_M-\mathcal{I}_k}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\|}_{\text{Fixed with }k}. \end{aligned} \tag{22}$$

In addition, $\mathbf{X}_i$ for $i\in\mathcal{I}_M-\mathcal{I}_k$ can be regarded as a zero matrix with (9). Thus, the second term on the RHS of (22) is simply σ_{k+1} because

$$\left\|\sum_{i\in\mathcal{I}_M-\mathcal{I}_k}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\|=\left\|\sum_{i\in\mathcal{I}_M-\mathcal{I}_k}\mathbf{B}_i\right\|=\sigma_{k+1}. \tag{23}$$

Then, with Lemma 1 and (23), (22) can be rewritten as

$$\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\|\le\sigma_1\delta. \tag{24}$$

The lower bound is a simple result of the Eckart–Young–Mirsky theorem [31], [32]. According to the theorem,

$$\sigma_{k+1}=\left\|\mathbf{B}-\sum_{i\in\mathcal{I}_k}\mathbf{B}_i\right\|\le\left\|\mathbf{B}-\widetilde{\mathbf{W}}\right\| \tag{25}$$

where the summation of $\mathbf{B}_i$ can be regarded as a rank-k approximation of $\mathbf{B}$ and $\widetilde{\mathbf{W}}\in\mathbb{C}^{N\times M}$ is an arbitrary matrix having with a rank of at most k. The lower bound is obviously valid because $\mathbf{A}\overline{\mathbf{X}}$ has a rank of at most k. □

One thing to note is that the RHS of (22) can be divided into two parts. The first part is controllable with tolerance δ_i. The second term is fixed when k is determined in (7). With this view, the lower bound (21) can also be justified. Also, the relative error ϵ_r approaches the lower bounds whenever (8) is solved more accurately. In other words, it should be emphasized that no matter how precisely the singular vectors are solved, even direct solver, the accuracy of the solution is limited by the lower bounds. Thus, for engineering applications, the tolerance is needed to be optimized. Consequently, Algorithm 1 proposes an optimal error tolerance in (10) to minimize the computational resources while also ensuring the bounds of the relative error of the target.

Even if the error of the matrix equation ϵ_r is bounded, one can argue that ϵ_r does not represent the error for each incident vector. One possible upper bound for all incident vectors is

$$\|\mathbf{b}_i-\mathbf{A}\overline{\mathbf{x}}_i\|\le\sigma_1\delta,\ \ (\forall i\in\mathcal{I}_M) \tag{26}$$

due to the definition of (1). This fact can be easily proven with (24) and a canonical basis vector, e.g., $\hat{e}_2=[0,1,\ldots,0]^T$. Note that the ith column vector of any matrix $\mathbf{W}\in\mathbb{C}^{N\times M}$ can be extracted by multiplication of $\hat{e}_i\in\mathbb{C}^M$ to $\mathbf{W}$. However, this approach is not tight enough for practical usage, especially when the compression rate is high (in that, k/M is small). Alternatively, the trend can be expected with root-mean-square (rms) errors for all incident vectors. Here, the upper and lower bounds of the rms errors are introduced by Corollary 1.

Corollary 1: Let (8) be solved with δ_i defined in (10) and set (9). Then,

$$\frac{\sigma_{k+1}}{\sqrt{M}}\le\epsilon_{\text{rms}}\le\frac{\sqrt{k}(\sigma_1\delta-\sigma_{k+1})+\sqrt{\sum_{i\in\mathcal{I}_M-\mathcal{I}_k}\sigma_i^2}}{\sqrt{M}} \tag{27}$$

where ϵ_{rms} is a quadratic mean of relative errors defined by

$$\epsilon_{\text{rms}}:=\sqrt{\frac{\sum_{i\in\mathcal{I}_M}\|\mathbf{b}_i-\mathbf{A}\overline{\mathbf{x}}_i\|^2}{M}}. \tag{28}$$

Proof: For an arbitrary matrix, $\mathbf{W}\in\mathbb{C}^{N\times M}$, it is easy to prove that the rms of the norms of the column vectors can be rewritten by the Frobenius norm divided by $\sqrt{M}$. Namely,

$$\epsilon_{rms}=\sqrt{\frac{\sum_{i\in\mathcal{I}_M}\|\mathbf{b}_i-\mathbf{A}\overline{\mathbf{x}}_i\|^2}{M}}=\frac{\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\|_F}{\sqrt{M}}. \tag{29}$$

The lower bound is easy to prove according to the properties of the matrix norm [23], [24]. Namely,

$$\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\|\le\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\|_F. \tag{30}$$

Using the lower bound of (21), (30) can be rewritten as

$$\frac{\sigma_{k+1}}{\sqrt{M}}\le\frac{\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\|}{\sqrt{M}}\le\epsilon_{\text{rms}}. \tag{31}$$

The upper bound can be derived with the similar idea of the proof in Lemma 1

$$\begin{aligned}\left\|\mathbf{B}-\mathbf{A}\overline{\mathbf{X}}\right\|_F &= \left\|\sum_{i\in\mathcal{I}_M}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\|_F \\ &\le \left\|\sum_{i\in\mathcal{I}_k}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\|_F + \left\|\sum_{i\in\mathcal{I}_M-\mathcal{I}_k}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\|_F \\ &= \left\|\left[\boldsymbol{\xi}_1\ \cdots\ \boldsymbol{\xi}_k\right]\boldsymbol{\Sigma}_{1:k}\right\|_F + \left\|\boldsymbol{\Sigma}_{k+1:M}\right\|_F \\ &\le \left\|\left[\frac{\boldsymbol{\xi}_1}{\|\boldsymbol{\xi}_1\|}\ \cdots\ \frac{\boldsymbol{\xi}_k}{\|\boldsymbol{\xi}_k\|}\right]\right\|_F\sqrt{\sum_{i\in\mathcal{I}_k}\sigma_i^2\|\boldsymbol{\xi}_i\|^2} \\ &\quad+\sqrt{\sum_{i\in\mathcal{I}_M-\mathcal{I}_k}\sigma_i^2} \\ &\le \sqrt{k}(\sigma_1\delta-\sigma_{k+1})+\sqrt{\sum_{i\in\mathcal{I}_M-\mathcal{I}_k}\sigma_i^2}. \end{aligned} \tag{32}$$

The second equality is valid because the Frobenius norm is invariant under unitary operation. For example,

$$\begin{aligned}\left\|\sum_{i\in\mathcal{I}_k}(\mathbf{B}_i-\mathbf{A}\mathbf{X}_i)\right\|_F &= \left\|\boldsymbol{\Xi}\mathbf{V}_{1:k}^H\right\|_F \\ &= \mathrm{trace}\Big(\left(\boldsymbol{\Xi}\mathbf{V}_{1:k}^H\right)^H\boldsymbol{\Xi}\mathbf{V}_{1:k}^H\Big) \\ &= \mathrm{trace}\big(\mathbf{V}_{1:k}\boldsymbol{\Xi}^H\boldsymbol{\Xi}\mathbf{V}_{1:k}^H\big) \\ &= \mathrm{trace}\big(\mathbf{V}_{1:k}^H\mathbf{V}_{1:k}\boldsymbol{\Xi}^H\boldsymbol{\Xi}\big) \\ &= \mathrm{trace}\big(\boldsymbol{\Xi}^H\boldsymbol{\Xi}\big) = \|\boldsymbol{\Xi}\|_F \\ &= \left\|\left[\boldsymbol{\xi}_1\ \cdots\ \boldsymbol{\xi}_k\right]\boldsymbol{\Sigma}_{1:k}\right\|_F \end{aligned} \tag{33}$$

where $\boldsymbol{\Xi} := \left[\boldsymbol{\xi}_1\ \cdots\ \boldsymbol{\xi}_k\right]\boldsymbol{\Sigma}_{1:k}$. The last inequality is from $\|\boldsymbol{\xi}_i\| \le \delta_i$ in (17). Finally, the upper bounds of (27) can be proven by dividing (32) by $\sqrt{M}$. □

E. Acceleration With a Variant of GCR for Multiple RHSs

Even with the optimal choice of the number of RHSs, further speedup can be achieved with the block version of the Krylov subspace methods. A simple variant block GCR for multiple RHSs is suggested to validate the extra speedup, as shown in Algorithm 2. In the algorithm, $\alpha, \beta,$ and $\gamma \in \mathbb{C}$. In addition, all vectors $z, \mathbf{r},$ and $\mathbf{s} \in \mathbb{C}^N$. The superscripts and subscripts represent the indices for singular vectors and iterations, respectively. The notation $< \mathbf{x}, \mathbf{y} >$ represent the inner products. Basically, Algorithm 2 constructs the search vectors ($\mathbf{s}_j$ and $\mathbf{v}_j$) to update the numerical solutions (z_j^i) and residual vectors ($\mathbf{r}_j^i$). A common choice [33], [34] for the initial guesses and initial search vectors is zero vectors ($\mathbf{r_0^i} = \mathbf{u_0^i}$) and $\mathbf{s}_j = \mathbf{r}_{j-1}$, respectively. One difference is that one residual vector having the largest norm becomes the next search vector. Thus, the algorithm is allowed to conduct only one matrix–vector multiplication (MVM) for every iteration ($\mathbf{v}_j \leftarrow \mathbf{A}\mathbf{s}_j$). Because all solution and residual vectors are updated with the same Krylov subspace, the proposed block GCR can reduce the number of required MVMs.

The truncation of Krylov subspace is needed for the block GCR due to limited computational resources. We propose

Algorithm 2 Variant Block GCR for Multiple RHSs

Given: $\mathbf{A}^{N\times N}$, $\mathbf{u}_i$, and δ_i, $\forall i \in \mathcal{I}_k$, in (8) and (10)
Result: z_i, $\forall i \in \mathcal{I}_k$

Initialize $z_0^i = \mathbf{0}$ and $\mathbf{r}_0^i = \mathbf{u}^i$, $\forall i \in \mathcal{I}_k$
Initialize $\mathbf{v}_j = \mathbf{0}$ and $\mathbf{s}_j = \mathbf{0}$, $\forall j \in \mathcal{I}_{j_{max}}$
Set $\mathbf{s}_1 = \mathbf{r}_0^p$, where $p = \underset{i\in\mathcal{I}_k}{\mathrm{argmax}} \left\|\mathbf{r}_0^i\right\|$
$\mathrm{flag}_i \leftarrow$ **false**, $\forall i \in \mathcal{I}_k$
for $j \leftarrow 1$ **to** j_{max} **do**
  $\mathbf{v}_j \leftarrow \mathbf{A}\mathbf{s}_j$ % Only one MVM
  for $q \leftarrow 1$ **to** $j-1$ **do**
    $\alpha \leftarrow < \mathbf{v}_j, \mathbf{v}_q >$
    $\mathbf{v}_j \leftarrow \mathbf{v}_j - \alpha\mathbf{v}_q$
    $\mathbf{s}_j \leftarrow \mathbf{s}_j - \alpha\mathbf{s}_q$
  end for
  $\beta \leftarrow \left\|\mathbf{v}_j\right\|$
  $\mathbf{v}_j \leftarrow \mathbf{v}_j/\beta$
  $\mathbf{s}_j \leftarrow \mathbf{s}_j/\beta$
  for $i \leftarrow 1$ **to** k **do** % for all singular vectors
    if ($\mathrm{flag}_i ==$ **true**) **cycle**
    $\gamma \leftarrow < \mathbf{r}_{j-1}^i, \mathbf{v}_j >$
    $z_j^i \leftarrow z_{j-1}^i + \gamma\,\mathbf{r}_j$
    $\mathbf{r}_j^i \leftarrow \mathbf{r}_{j-1}^i - \gamma\,\mathbf{v}_j$
    if ($\left\|\mathbf{r}_j^i\right\| \le \delta_i$) $\mathrm{flag}_i \leftarrow$ **true**
  end for
  if ($\mathrm{flag}_i ==$ **true**, $\forall i \in \mathcal{I}_k$) **exit**
  $\mathbf{s}_{j+1} = \mathbf{r}_j^p$, where $p = \underset{i}{\mathrm{argmax}}\left\|\mathbf{r}_j^i\right\|$
end for
$z_i \leftarrow z_j^i$

that one search vector can be abandoned and replaced with a new search vector every iteration after reaching the maximum number of stored Krylov vectors. The replacement vector is chosen with the following index:

$$q = \underset{l}{\mathrm{argmin}} < \mathbf{v}_l, \mathbf{v}_j > \tag{34}$$

where $\mathbf{v}_j$ is the current search vector and $\mathbf{v}_l$ are the stored previous search vectors. The vector $\mathbf{v}_q$ is the most similar to the new search vector. The effect of the block GCR with truncation is shown in Section III-D.

III. Numerical Validation

This section gives the numerical evidence used to validate Theorem 1 and Corollary 1. We computed and compared the upper and lower bounds and the actual relative error ϵ_r with different tolerances. In addition, the numerical examples indicate the efficiency of the computation of angular responses with the proposed algorithm. Throughout the experiments, we applied the algorithm to the discontinuous Galerkin integral equation (IEDG) method [35] with the multilevel fast multipole method (MLFMM) [36]. The targets were discretized by nonconformal and mixed triangles and quadrilaterals. All computations were conducted using workstations in the Owens Cluster at the Ohio Supercomputer Center (OSC) [37].

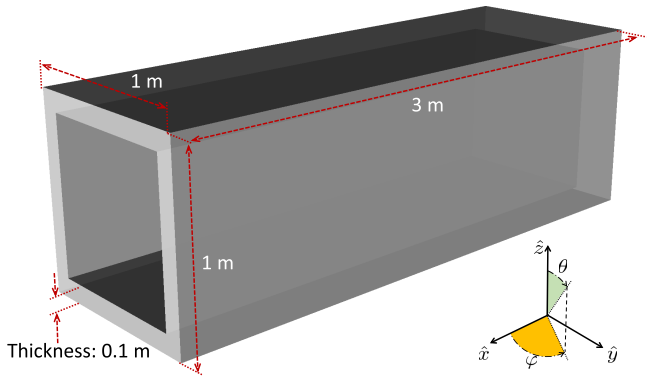


Fig. 1. PEC cuboid cavity.

The first two targets were computed with nodes with two 14-core Intel Xeon E5-2680 v4 (2.40 GHz) processors and 128 GB memory. The last fighter jet example was computed with nodes with four 12-core Intel Xeon E5-4830 v3 (2.10 GHz) processors and 1536 GB memory. The three targets defined in this section have different electrical sizes, geometrical complexities, and materials.

A. Target

1) Perfect Electric Conductor Cuboid Cavity: In Fig. 1, the first example is plotted. The target is a perfect electric conductor (PEC) cuboid cavity with an open left side. The frequency is 600 MHz, and the size of the cube that surrounds the cavity is $6\lambda_0$, where λ_0 is the free-space wavelength. The mesh is prepared with size $\lambda_0/6$, and the number of unknowns is 30 704. The desired angular responses are $-180° \leq \varphi \leq 180°$ and $\theta = 90°$, with fixed polarization angles 0°. We set the number of incident angles M as 181 with $\Delta\varphi = 2°$ and $\Delta\theta = 0$, whose values meet (4). The tolerance ϵ_r is set to the values ranging from 10^{-7} to 10^{-3}.

2) Finite-Conductivity Cylindrical Cavity: The next example is more difficult to solve; it is a cylindrical cavity made of real metal, as described in Fig. 2. This target is a high-resonance structure with thin-aperture slots located on the front of the cylinder (0.508 × 50.8 mm). The height is 609.6 mm, and the inner radius is 101.6 mm. The thickness of the metal is 6.35 mm. The metal has a finite conductivity, $\sigma = 2.6 \times 10^7$ [S/m], so we apply the impedance boundary condition (IBC) [38], [39], [40], [41], [42] to approximate the imperfect electric conductivity. The first theoretical resonance frequency is at 1129.391 MHz for transverse magnetic (TM) mode. In this experiment, the operating frequency is 1132.4207 MHz, which is the experimental resonance frequency with the given mesh and simulation setup. The cylinder is discretized with a maximum mesh size of approximately $\lambda_0/15$, and the number of unknowns is 28 944. In addition, the size of the cube enclosing the cylinder in wavelengths is approximately $2.35\lambda_0$. For this example, the 2-D angular responses are computed. The interesting angular sector is $-60° \leq \varphi \leq 60°$ and $70° \leq \theta \leq 110°$ with a 0° polarization angle. With (4), we set $\Delta\varphi = 4°$ and $\Delta\theta = 4°$ for the azimuthal and elevation angles, respectively. Thus, the total number is $M = 341$. The tolerance ϵ_r also ranges from 10^{-7} to 10^{-3}. With this target, we can show the effectiveness of the proposed method in the case of a 2-D angular sweep and targets made from a realistic material.

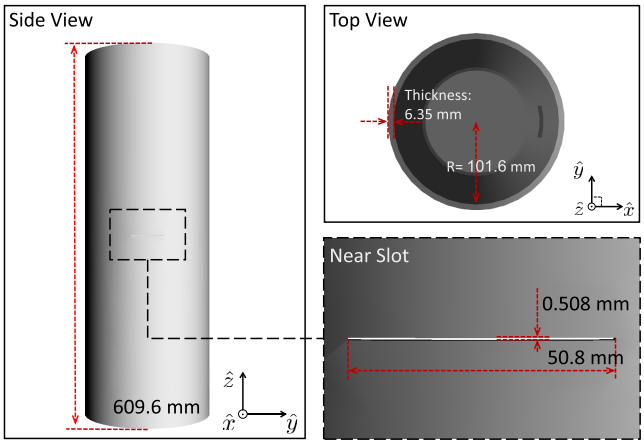


Fig. 2. Finite-conductivity cylindrical cavity.

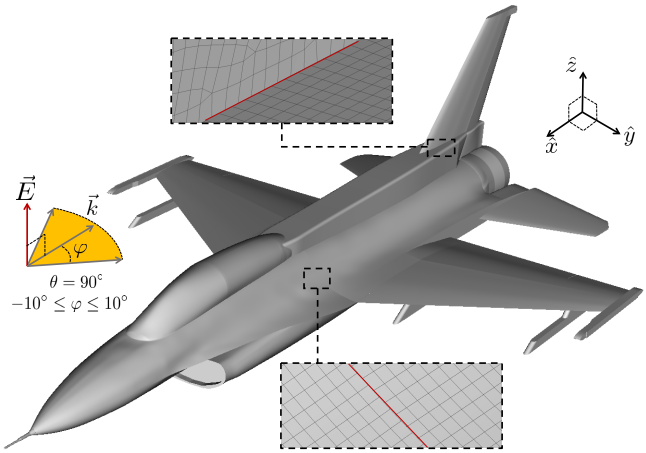


Fig. 3. Problem domain of the PEC F-16 fighter jet and examples of nonconformal meshes.

3) PEC F-16 Fighter Jet: To show the robustness of the error control of the proposed algorithm, we choose an electrically large target, the PEC F-16 fighter jet (Figs. 3 and 4), which has a large electrical size and a complicated geometry. Taking advantage of IEDG, all targets can be divided into 45 parts and meshed independently. As a result, a nonconformal mesh is prepared, and the total number of unknowns is 7 700 888, with an average mesh size of $\lambda_0/5$.

Note that the intake is closed with the flat surface in this example. In this study, the operating frequency is 5 GHz; accordingly, the longest dimension of the target is approximately $250.33\lambda_0$. In addition, the desired angle of the incident fields is fixed at $\theta = 90°$, and $-10° \leq \varphi \leq 10°$ is swept, with a 0° polarization angle. The optimal angular step we choose is 0.1° according to (4); therefore, a total of 201 incident vectors are used. The tolerance ϵ_r varies from 10^{-4} to 10^{-2}.

B. Singular Values

Fig. 5 shows the singular values for the numerical examples on a semilogarithmic scale. The singular values for

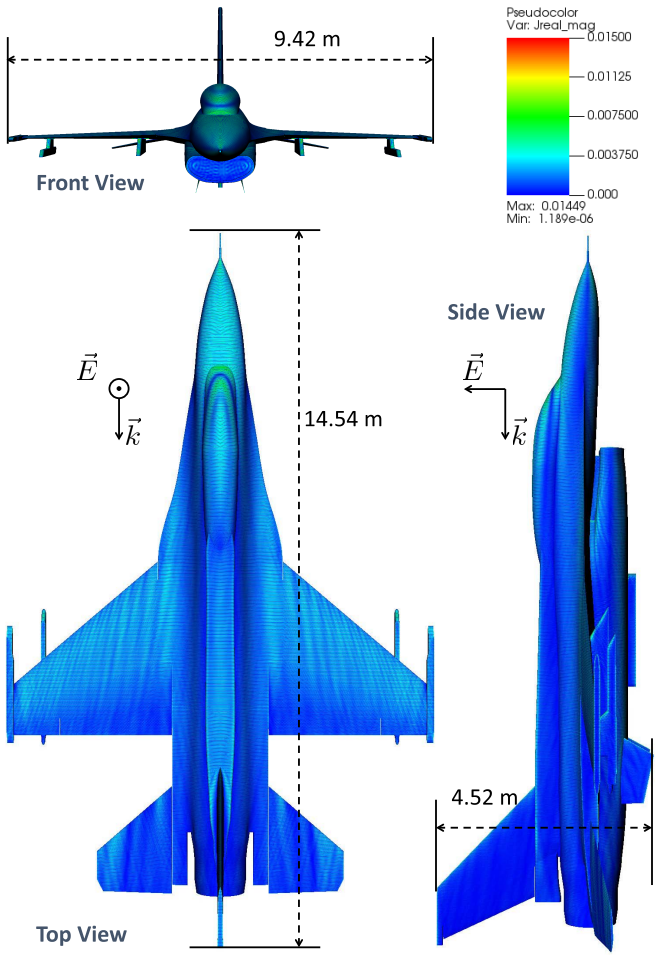


Fig. 4. Surface electric current distribution of the PEC F-16 fighter jet at a head-on incident angle ($\varphi = 0°$ and $\theta = 90°$).

TABLE I
NUMBER OF SINGULAR VECTORS, k, AS A FUNCTION OF TOLERANCE δ

Target	δ					
	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
PEC Cuboid	–	28	31	34	37	39
IBC Cylinder	–	34	49	63	79	96
PEC F-16	61	64	67	–	–	–

the excitation matrix $\mathbf{B}$ are plotted with solid lines. The other lines represent the singular values of the error matrices ($\mathbf{E} := \mathbf{B} - \mathbf{A}\overline{\mathbf{X}}$) with respect to the target tolerance δ. It should be noted that the computation of $\sigma(\mathbf{E})$ is just for debugging purposes, to verify the error bounds. The maximum singular values of incident matrices $\mathbf{B}$ and $\mathbf{E}$ are the normalization factor and the actual error in Theorem 1, respectively. In addition, Table I presents the number of singular vectors k, which is determined by (7), in terms of the given tolerance δ.

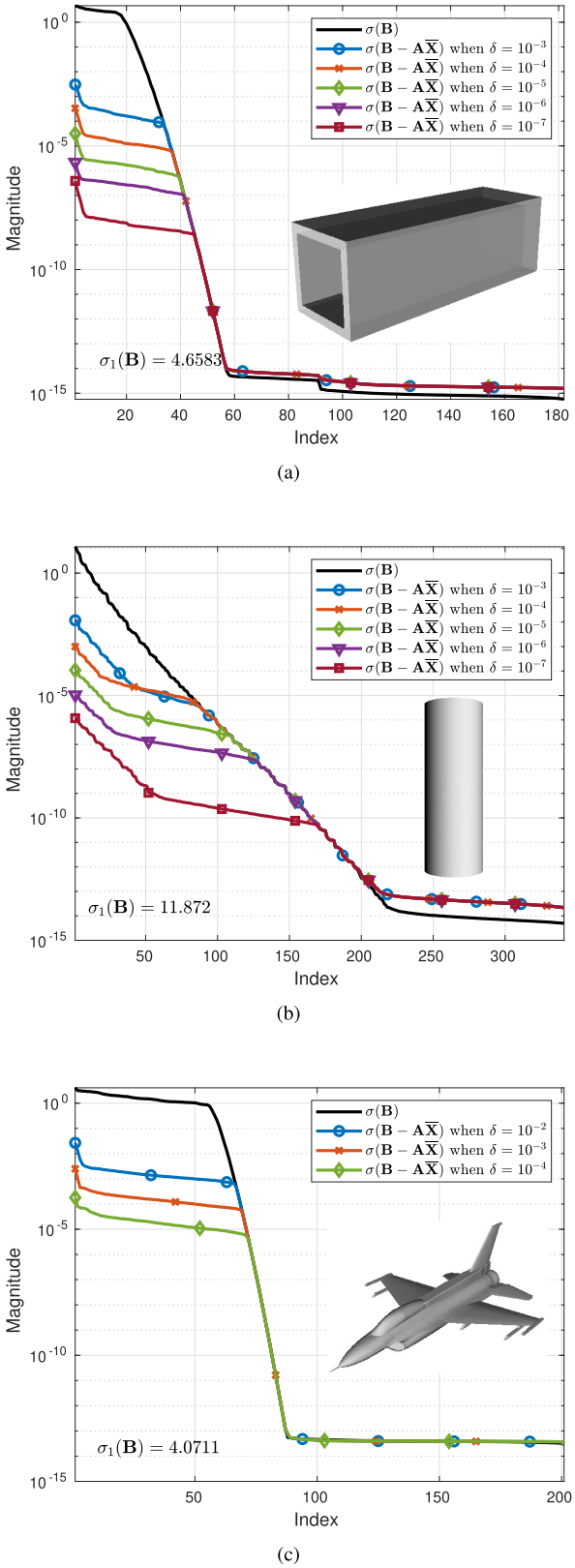
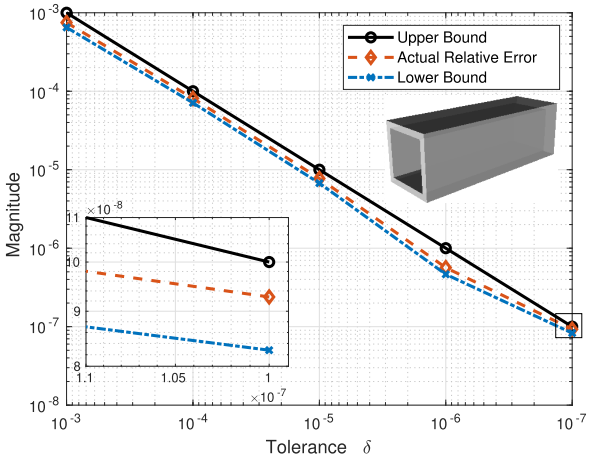


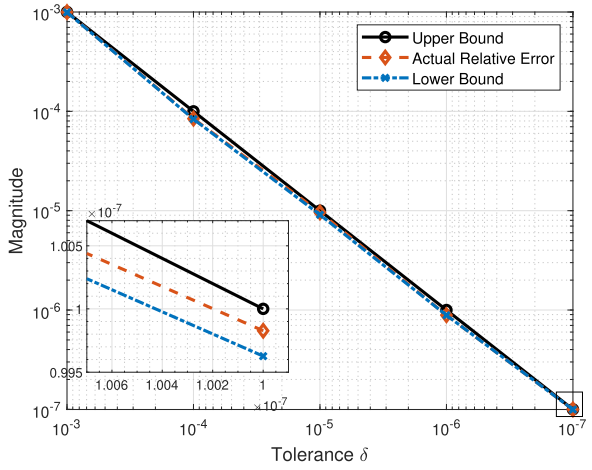
Fig. 5. Singular value distribution (σ) for various targets. (a) PEC cuboid cavity. (b) IBC cylinder cavity. (c) PEC F-16 fighter jet.

C. Error Bounds

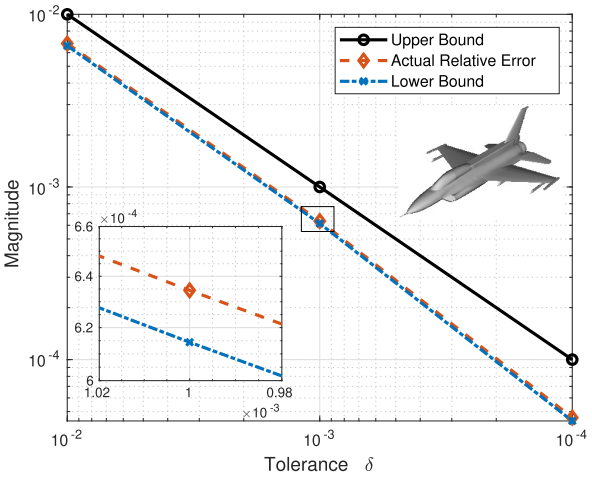
As shown in Fig. 6, the error bounds of Theorem 1 are numerically validated. In Fig. 6, the actual error ϵ_r and the theoretical bounds are provided with the given tolerances on logarithmic scales. As mentioned above, the actual relative errors are computed with the singular values of $\mathbf{E}$, which are shown by the red line with the diamond marker. The black



(a)

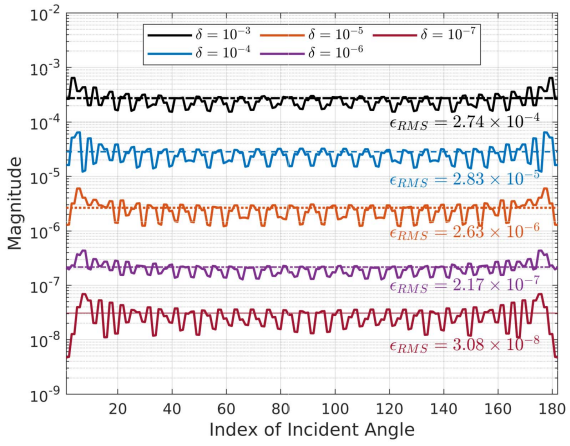


(b)

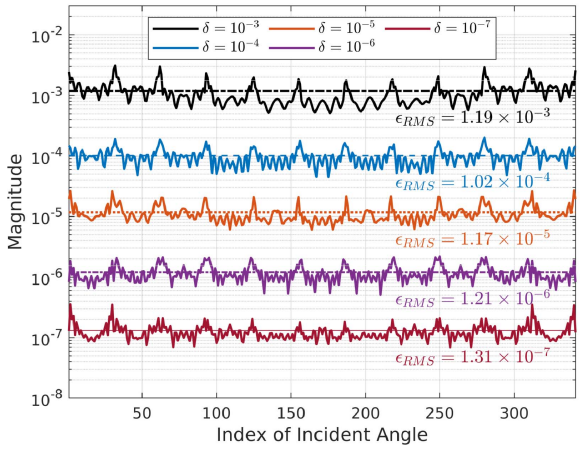


(c)

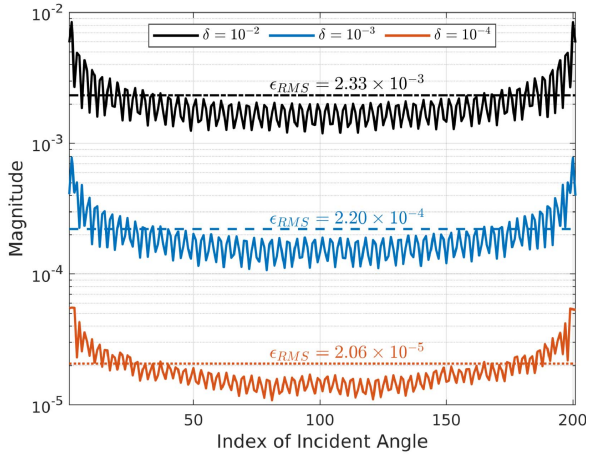
Fig. 6. Actual relative error (ϵ_r) and its bounds with respect to the given tolerance δ. (a) PEC cuboid cavity. (b) IBC cylinder cavity. (c) PEC F-16 fighter jet.



(a)



(b)



(c)

Fig. 7. Individual error ($\|\mathbf{b}_i - \mathbf{A}\overline{\mathbf{x}}_i\|$) for all incident angles. (a) PEC cuboid cavity. (b) IBC cylinder cavity. (c) PEC F-16 fighter jet.

solid line represents the upper bound, which is identical to the predetermined tolerance δ. The blue lines with asterisk markers represent the lower bounds. The magnified subfigures are depicted at the bottom left of the figures. According to the results, it is numerically confirmed that the actual errors reside between the error bounds.

In addition, in Fig. 7, the semilogarithmic plots provide the individual errors of all incident vectors and their rms values in terms of the global tolerance δ. Fig. 8 shows the rms values with the theoretical upper and lower bounds as a function of

the global tolerance δ on logarithmic scales. All errors and rms values are obtained by actual computation of $\mathbf{E}$ from the recovered solution $\overline{\mathbf{X}}$. The actual rms error, the upper bounds, and the lower bounds are shown in a similar way to that of Fig. 6. As shown in Fig. 8, the numerical examples satisfy Corollary 1.

As shown in this section, both the actual ϵ_r and ϵ_{rms} are bounded, as proven in Theorem 1 and Corollary 1. Hence, our suggested approach can readily estimate and control the error of the solutions with a predetermined global tolerance. Section III-D numerically validates the benefit of the proposed algorithm, i.e., different tolerances for different singular vectors, as well as the block version of the GCR.

D. Computational Costs

In this section, the computational costs are provided by comparing the wall time, the number of MVMs, and the required memory for the iterative solver. The experiment is designed with the following four cases. The first case (Case 1) is regarded as a baseline with the conventional way to solve with multiple RHSs. Namely, after assembling the system matrix, each incident vector is solved individually. To make a fair comparison, the tolerances for each incident vector are taken from the actual data of $\|\mathbf{b}_i - \mathbf{A}\overline{\mathbf{x}}_i\|$ in Fig. 7. For the PEC F-16 target, we estimate the number of MVMs and the wall time for the solution process using the 20 angles data samples due to the observation that the number of iterations is nearly identical for all incident angles. The next case (Case 2) uses a rank-k approximation of $\mathbf{B}$ in (7), but the tolerance of each singular vector is set to

$$\delta_i = \frac{1}{\sqrt{k}}\left(\delta - \frac{\sigma_{k+1}}{\sigma_1}\right) \tag{35}$$

for $i \in \mathcal{I}_k$. This is a reasonable choice to ensure that the accuracy between the second and third cases is similar because

$$\left\|\mathbf{B} - \mathbf{A}\overline{\mathbf{X}}\right\| \le \sqrt{k}\max_{i\in\mathcal{I}_k}(\sigma_i\delta_i) + \sigma_{k+1} = \sigma_1\delta \tag{36}$$

from (19) and (22). In addition, we use Algorithm 1 for the third (Case 3) and the last cases (Case 4). The conventional GCR and the modified block GCR introduced in Algorithm 2 are used for the third and fourth cases, respectively. Note that the data discussed in Sections III-B and III-C were collected from the third case. Also, the number of truncations of the Krylov subspace for both the conventional GCR and the modified block GCR for MRHSs (j_{max}) is fixed 100. Fig. 9 shows the number of MVMs with respect to the target tolerances for each case. By comparing Cases 1 and 2 in Fig. 9, we can observe a significant reduction of the number of RHSs with a low-rank approximation as expected. According to Case 3, the recommended tolerance (δ_i) for each singular vector in (10) improves the number of MVMs compared to Case 2. In addition, the results from comparing Cases 3 and 4 further show the meaningful reduction in the number of MVMs. Table II shows the wall time statistics of the above experiment. As expected, the solution time for all examples tends to be proportional to the number of MVMs because the majority of the time-consuming portion is MVM during the iterative solution process. Note that the wall time in Table II does not include the SVD times that are 20.3 s, 27.8 s, and 39.3 min for targets 1, 2, and 3, respectively.

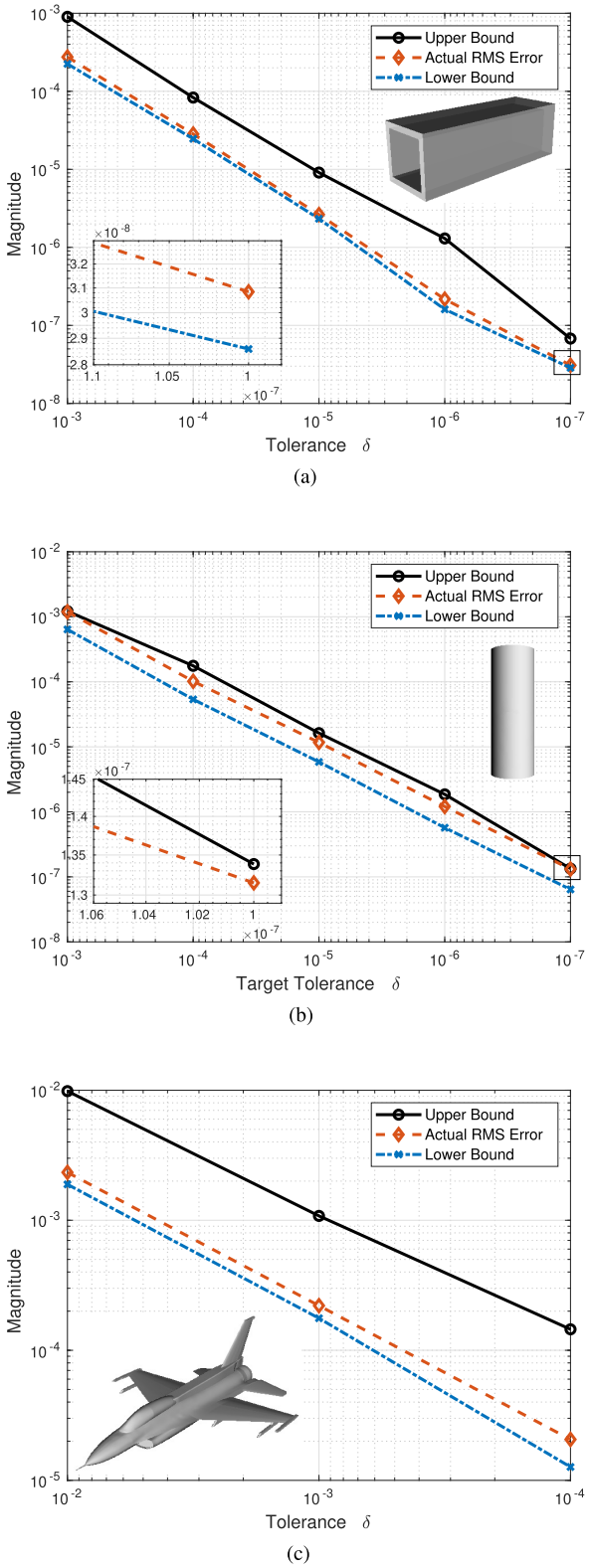


Fig. 8. Actual rms of the individual error (ϵ_{rms}) and its bounds with respect to the given tolerance δ. (a) PEC cuboid cavity. (b) IBC cylinder cavity. (c) PEC F-16 fighter jet.

Table III shows the summary of the memory requirements during the solution process. The memory for the 100 search

TABLE II
SOLUTION TIME (WALL TIME) AS A FUNCTION OF TOLERANCE δ

Example	δ	Case 1	Case 2	Case 3	Case 4
PEC Cuboid** (III-A1)	10^{-3}	44 : 05	09 : 27	07 : 50	01 : 14
	10^{-4}	57 : 50	11 : 14	09 : 22	01 : 27
	10^{-5}	63 : 19	13 : 03	11 : 12	01 : 42
	10^{-6}	67 : 04	14 : 43	12 : 24	01 : 53
	10^{-7}	73 : 38	16 : 16	13 : 46	02 : 12
IBC Cylinder** (III-A2)	10^{-3}	19 : 07	05 : 10	04 : 01	01 : 01
	10^{-4}	41 : 53	07 : 22	04 : 18	01 : 14
	10^{-5}	48 : 09	10 : 02	06 : 52	01 : 32
	10^{-6}	50 : 59	13 : 05	09 : 06	01 : 53
	10^{-7}	53 : 07	17 : 26	14 : 32	03 : 17
PEC F-16*** (III-A3)	10^{-2}	96 : 47*	37 : 12	29 : 25	26 : 15
	10^{-3}	182 : 04*	63 : 00	48 : 58	40 : 55
	10^{-4}	318 : 43*	94 : 34	72 : 36	54 : 20

* Estimated Value
** Wall time denotes in minutes and seconds (mm:ss)
*** Wall time denotes in hours and minutes (hh:mm)

TABLE III
PEAK MEMORY FOR SOLUTION PROCESS AS A FUNCTION OF TOLERANCE δ

Example	δ	Case 1	Case 2	Case 3	Case 4
PEC Cuboid** (III-A1)	10^{-3}	180.51	108.83	108.83	121.95
	10^{-4}		110.23	110.23	124.75
	10^{-5}		111.64	111.64	127.57
	10^{-6}		113.04	113.04	130.38
	10^{-7}		113.98	113.98	132.25
IBC Cylinder** (III-A2)	10^{-3}	241.98	106.40	106.40	121.41
	10^{-4}		113.02	113.02	124.66
	10^{-5}		119.20	119.20	147.03
	10^{-6}		126.27	126.27	161.16
	10^{-7}		133.78	133.78	176.18
PEC F-16*** (III-A3)	10^{-2}	47.02*	30.95	30.95	37.95
	10^{-3}		31.29	31.29	38.64
	10^{-4}		31.64	31.64	39.33

* Estimated Value
** Megabytes
*** Gigabytes

vectors (95.71 MB, 91.38 MB, and 23.95 GB for targets 1, 2, and 3, respectively) is embedded in the numbers in Table III. It requires for Case 1 to store fixed 181, 341, and 201 RHS vectors for different tolerances. On the other hand, Cases 2–4 need the different memories depending on the number of singular vectors, as shown in Table I. Also, Cases 2 and 3 have identical values since both cases use exactly the same procedure except for the tolerance of the iterative solver (δ_i). Case 4 requires the additional memory for the solution and residual vectors (z^i and $\mathbf{r}^i$) in Algorithm 2 for all singular vectors. Note that the values in Table III are minor compared to the memory requirements of the system matrices ($\mathbf{A}$), which are 3.823, 7.293, and 478.5 GB for each target.

E. Application: Monostatic Radar Cross Section of F-16

In this section, a practical but challenging example is considered, the computation of the 2-D monostatic radar cross section (RCS) of an F-16 with a deep cavity intake. Basically, the problem geometry is similar to that of the PEC F-16 in Section III-A3; however, there are three differences. First, the operating frequency is 8 GHz, which is the typical starting

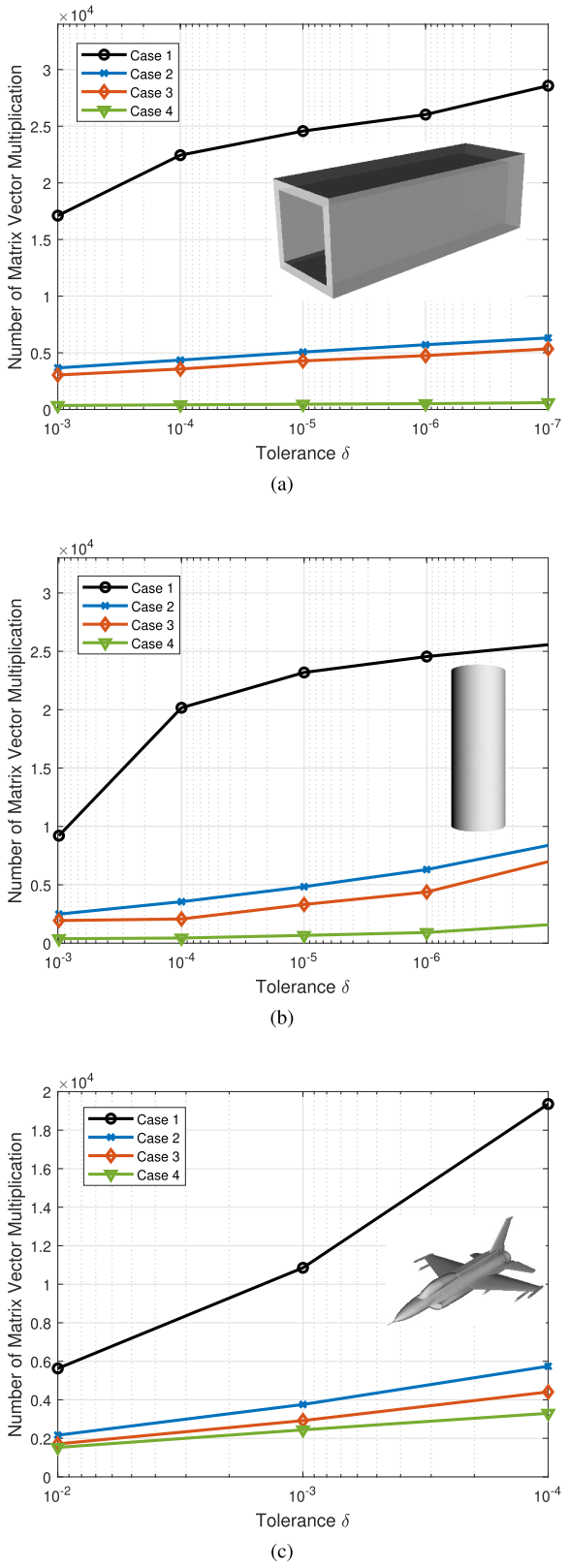


Fig. 9. Number of MVM for each case. (a) PEC cuboid cavity. (b) IBC cylinder cavity. (c) PEC F-16 jet fighter.

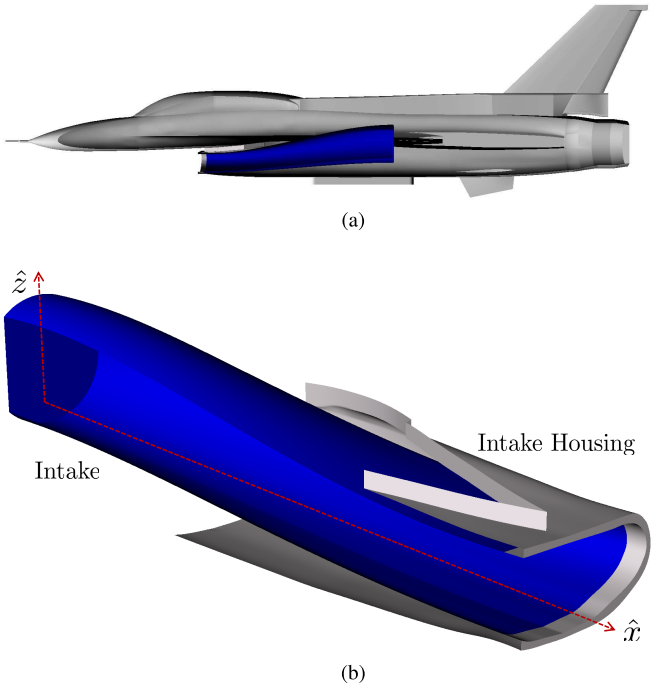


Fig. 10. PEC F-16 model with a deep cavity intake (blue parts). (a) Side view. (b) Cross section at the intake of the F-16 jet fighter.

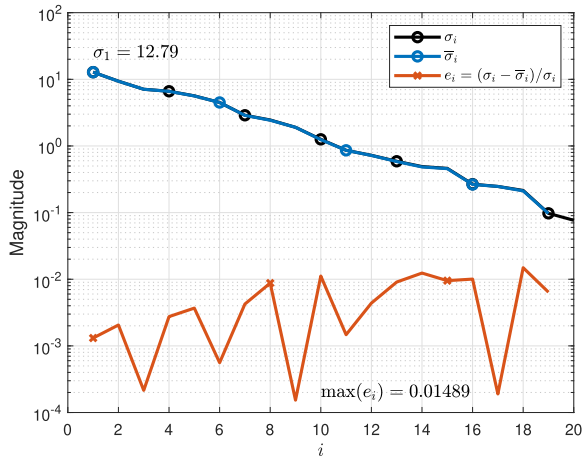


Fig. 11. Singular value distribution for the first block for the F-16 model with a deep cavity intake ($-2.0° \le \varphi \le -1.05°$ and $88.0° \le \theta \le 89.0°$).

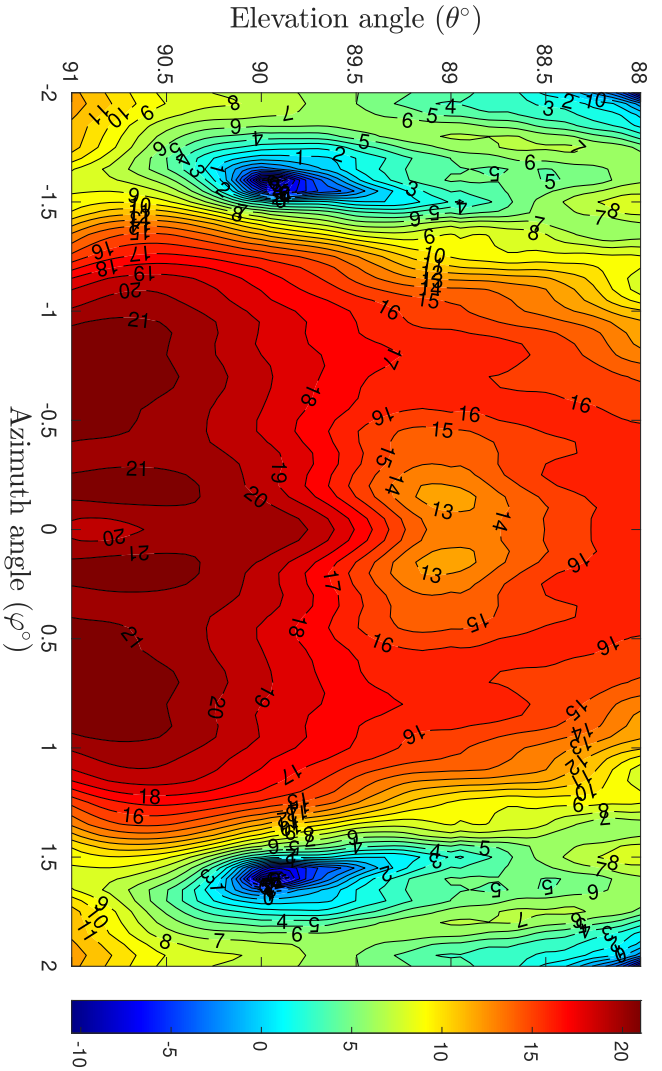


Fig. 12. Monostatic RCS (unit: dBsm) of the F-16 model with a deep cavity intake.

frequency of the X-band [43]. With the higher frequency, the target size is approximately 400.53λ; hence, the number of basis functions is 23 194 584 with $\lambda/5$ meshes. Second, different from the closed intake in Section III-A3, the engine intake is modeled as a deep concave structure, which is the resonance structure shown in Fig. 10. One can expect that the convergence behavior is deteriorated due to the resonance structure. Finally, the required number of angular responses is significantly larger according to (4). In our experiment, $\Delta\varphi$ and $\Delta\theta$ are both 0.05°. The angular responses of interest are $-2° \le \varphi \le 2°$ and $88° \le \theta \le 91°$ with a 0° polarization angle. As a result, the total number of RHS is $M = 5001$. Because of the batch limits at the OSC, the 5001 RHSs were divided into 12 blocks; hence, each block typically consisted of 420 RHSs.

Since we observed that each block behaves similarly, the data of the first block (a typical block) are presented as follows. The incident angles for the first block are $-2.0° \le \varphi \le -1.05°$ and $88.0° \le \theta \le 89°$. Fig. 11 shows the distribution of the first 20 singular values of the block. According to Algorithm 1, the first 18 singular vectors were retained to solve from 420 RHS vectors in this block. In addition, the singular values obtained by the conventional SVD (σ_i) and PCA ($\overline{\sigma}_i$) are plotted in this figure. Furthermore, the relative difference, $e_i := (\sigma_i - \overline{\sigma}_i)/\sigma_i$, is presented in Fig. 11 to validate that the PCA approach is as accurate as the conventional SVD. Fig. 12 shows the monostatic RCS of PEC F-16 with a deep cavity intake. As a self-consistency, the RCS plot is symmetric with respect to $\phi = 0°$. One interesting point is that the peak points are not located at the head-on incident direction ($\varphi = 0°$ and $\theta = 90°$) but slightly lower and to the side of it ($\varphi = 21.7°$ and $\theta = 90.75°$). Finally, Table IV summarizes the wall time for the entire 12 blocks (5001 RHS vectors). In this table, the "naive approach" is quite similar to case 1 in Section III-D, but the tolerance is fixed to 10^{-2}. In addition, the total solution time is estimated by the solution time for the first 20 RHS

TABLE IV
WALL TIME COMPARISON (HH)

Target	Naive approach	Approach in this paper	
Matrix assembly time	6.45	6.45	
Solution time	$23,249^{*}$	316.23	
Factorization time	–	98.4 (SVD)*	7.86 (PCA)

* Estimated Value

vectors. On the other hand, the "approach in this article" used Algorithms 1 and 2, which is the same as Case 4 in Section III-D. As shown in Table IV, the speedup is almost 73.6 times if PCA is utilized. For both approaches, the peak memory is the same, at 1110.35 GB, because the majority of the memory required is for the system matrix (1041.22 GB) and 100 Krylov vectors (69.13 GB). The comparison between the proposed factorization using the conventional SVD and using PCA is also included in Table IV. Wall time for PCA for the entire 5001 RHSs is about 7.86 h, which is a significant improvement compared to 98.4 h. Therefore, the speedup of the factorization process is about 12.51. Note that the factorization time for the entire 12 blocks with conventional SVD is estimated based on 8.2 h of the first (typical) block of 420 RHS vectors.

IV. CONCLUSION

In this article, an efficient and error-controllable algorithm has been proposed. Also, it was proved and numerically validated that the numerical error is bounded with the algorithm. In addition, the block version of GCR can effectively accelerate the solution time with a limited increase in memory. Although the proposed method was applied to electromagnetic scattering via the IEDG framework, the algorithm can be extended to any other method with multiple RHSs, for example, FEM-based algorithm [44], [45], [46], [47].

ACKNOWLEDGMENT

This article describes objective technical results and analysis. Any subjective views or opinions that might be expressed in this article do not necessarily represent the views of the U.S. Department of Energy or the United States Government. The authors would like to thank their colleague, Dr. Haobo Yuan, National Key Laboratory of Antennas and Microwave Technology, Xi'an, China, for useful discussions.

REFERENCES

[1] Y. Liu, Y. R. Mao, Y. J. Xie, and Z. H. Tian, "Evaluation of passive intermodulation using full-wave frequency-domain method with nonlinear circuit model," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5754–5757, Jul. 2016.

[2] R. Hong, K. Chen, X. Hou, Q. Sun, N. Liu, and Q. H. Liu, "Mixed finite element method for full-wave simulation of bioelectromagnetism from DC to microwave frequencies," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 10, pp. 2765–2772, Oct. 2020.

[3] Y. Ren, Y. Chen, Q. Zhan, J. Niu, and Q. H. Liu, "A higher order hybrid SIE/FEM/SEM method for the flexible electromagnetic simulation in layered medium," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2563–2574, May 2017.

[4] G. W. Stewart, *Matrix Algorithms: Basic Decompositions*, vol. 1. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, Jan. 1998, doi: 10.1137/1.9781611971408.

[5] W. Hackbusch, "A sparse matrix arithmetic based on H-matrices. Part I: Introduction to H-matrices," *Computing*, vol. 62, no. 2, pp. 89–108, 1999, doi: 10.1007/s006070050015.

[6] B. Zhou and D. Jiao, "Linear-complexity direct finite element solver accelerated for many right hand sides," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APSURSI)*, Jul. 2014, pp. 1383–1384.

[7] 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.

[8] V. Simoncini and E. Gallopoulos, "Convergence properties of block GMRES and matrix polynomials," *Linear Algebra Appl.*, vol. 247, pp. 97–119, Nov. 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0024379595000933

[9] J. Meng, P.-Y. Zhu, and H.-B. Li, "A block GCROT(m, k) method for linear systems with multiple right-hand sides," *J. Comput. Appl. Math.*, vol. 255, pp. 544–554, Jan. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377042713003129

[10] M. L. Parks, E. De Sturler, G. Mackey, D. D. Johnson, and S. Maiti, "Recycling Krylov subspaces for sequences of linear systems," *SIAM J. Sci. Comput.*, vol. 28, no. 5, pp. 1651–1674, Oct. 2006, doi: 10.1137/040607277.

[11] Z. Peng, M. B. Stephanson, and J.-F. Lee, "Fast computation of angular responses of large-scale three-dimensional electromagnetic wave scattering," *IEEE Trans. Antennas Propag.*, vol. 58, no. 9, pp. 3004–3012, Sep. 2010.

[12] N. V. Venkatarayalu, Y.-B. Gan, K. Zhao, and J.-F. Lee, "Fast monostatic RCS computation in FEM based solvers using QR decomposition," in *Proc. 1st Eur. Conf. Antennas Propag.*, Nov. 2006, pp. 1–5.

[13] B.-Y. Wu and X.-Q. Sheng, "Application of asymptotic waveform evaluation to hybrid FE-BI-MLFMA for fast RCS computation over a frequency band," *IEEE Trans. Antennas Propag.*, vol. 61, no. 5, pp. 2597–2604, May 2013.

[14] X. C. Wei, Y. J. Zhang, and E. P. Li, "The hybridization of fast multipole method with asymptotic waveform evaluation for the fast monostatic RCS computation," *IEEE Trans. Antennas Propag.*, vol. 52, no. 2, pp. 605–607, Feb. 2004.

[15] E. K. Miller, "Model-based parameter estimation in electromagnetics. III. Applications to EM integral equations," *IEEE Antennas Propag. Mag.*, vol. 40, no. 3, pp. 49–66, Jun. 1998.

[16] N. Mutonkole, E. R. Samuel, D. I. L. D. Villiers, and T. Dhaene, "Parametric modeling of radiation patterns and scattering parameters of antennas," *IEEE Trans. Antennas Propag.*, vol. 64, no. 3, pp. 1023–1031, Mar. 2016.

[17] M. Kazempour and L. Gurel, "Fast solution of electromagnetic scattering problems with multiple excitations using the recompressed adaptive cross approximation," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APSURSI)*, Jul. 2014, pp. 745–746.

[18] A. Schroder, H. D. Bruns, and C. Schuster, "A hybrid approach for rapid computation of two-dimensional monostatic radar cross section problems with the multilevel fast multipole algorithm," *IEEE Trans. Antennas Propag.*, vol. 60, no. 12, pp. 6058–6061, Dec. 2012.

[19] X.-M. Pan and X.-Q. Sheng, "Accurate and efficient evaluation of spatial electromagnetic responses of large scale targets," *IEEE Trans. Antennas Propag.*, vol. 62, no. 9, pp. 4746–4753, Sep. 2014.

[20] M. S. Chen, F. L. Liu, H. M. Du, and X. L. Wu, "Compressive sensing for fast analysis of wide-angle monostatic scattering problems," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 1243–1246, 2011.

[21] S.-R. Chai and L.-X. Guo, "Compressive sensing for monostatic scattering from 3-D NURBS geometries," *IEEE Trans. Antennas Propag.*, vol. 64, no. 8, pp. 3545–3553, Aug. 2016.

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

[23] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[24] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2013.

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

[26] M. Nilsson, "Rapid solution of parameter-dependent linear systems for electromagnetic problems in the frequency domain," *IEEE Trans. Antennas Propag.*, vol. 53, no. 2, pp. 777–784, Feb. 2005.
[27] S. Ubaru and Y. Saad, "Fast methods for estimating the numerical rank of large matrices," in *Proc. 33rd Int. Conf. Mach. Learn.*, vol. 48, M. F. Balcan and K. Q. Weinberger, Eds. New York, NY, USA, Jun. 2016, pp. 468–477. [Online]. Available: https://proceedings.mlr.press/v48/ubaru16.html
[28] V. Rokhlin, A. Szlam, and M. Tygert, "A randomized algorithm for principal component analysis," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 3, pp. 1100–1124, 2009, doi: 10.1137/080736417.
[29] 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. [Online]. Available: https://www.pnas.org/content/104/51/20167
[30] F. Woolfe, E. Liberty, V. Rokhlin, and M. Tygert, "A fast randomized algorithm for the approximation of matrices," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 3, pp. 335–366, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1063520307001364
[31] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936, doi: 10.1007/BF02288367.
[32] L. Mirsky, "Symmetric gauge functions and unitarily invariant norms," *Quart. J. Math.*, vol. 11, no. 1, pp. 50–59, 1960, doi: 10.1093/qmath/11.1.50.
[33] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, "Variational iterative methods for nonsymmetric systems of linear equations," *SIAM J. Numer. Anal.*, vol. 20, no. 2, pp. 345–357, Apr. 1983, doi: 10.1137/0720023.
[34] F. J. Lingen, "A generalised conjugate residual method for the solution of non-symmetric systems of equations with multiple right-hand sides," *Int. J. Numer. Methods Eng.*, vol. 44, no. 5, pp. 641–656, Feb. 1999. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-0207%2819990220%2944%3A5%3C641%3A%3AAID-NME520%3E3.0.CO%3B2-P
[35] 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, Jul. 2013.
[36] J. M. Song, C.-C. Lu, and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propag.*, vol. 45, no. 10, pp. 1488–1493, Oct. 1997.
[37] (2016). *Owens, Ohio Supercomputer Center*. [Online]. Available: http://osc.edu/ark:/19495/hpc6h5b1
[38] T. B. A. Senior, "Impedance boundary conditions for imperfectly conducting surfaces," *Appl. Sci. Res., B*, vol. 8, pp. 418–436, Dec. 1960, doi: 10.1007/BF02920074.
[39] L. Medgyesi-Mitschang and J. Putnam, "Integral equation formulations for imperfectly conducting scatterers," *IEEE Trans. Antennas Propag.*, vol. AP-33, no. 2, pp. 206–214, Feb. 1985.
[40] A. Bendali, M. B. Fares, and J. Gay, "A boundary-element solution of the Leontovitch problem," *IEEE Trans. Antennas Propag.*, vol. 47, no. 10, pp. 1597–1605, Oct. 1999.
[41] S. Yan and J.-M. Jin, "Self-dual integral equations for electromagnetic scattering from IBC objects," *IEEE Trans. Antennas Propag.*, vol. 61, no. 11, pp. 5533–5546, Nov. 2013.
[42] X.-W. Huang and X.-Q. Sheng, "A discontinuous Galerkin self-dual integral equation method for scattering from IBC objects," *IEEE Trans. Antennas Propag.*, vol. 67, no. 7, pp. 4708–4717, Jul. 2019.
[43] *IEEE Standard Letter Designations for Radar-Frequency Bands*, IEEE Standard 521–2002 (Revision IEEE Std 521-1984), pp. 1–10, 2003.
[44] P. Monk, *Finite Element Methods for Maxwell's Equations* (Numerical Mathematics and Scientific Computation). Oxford, U.K.: Oxford Univ. Press, 2003.
[45] J. F. Lee, D. K. Sun, and Z. J. Cendes, "Tangential vector finite elements for electromagnetic field computation," *IEEE Trans. Magn.*, vol. 27, no. 5, pp. 4032–4035, Sep. 1991.
[46] S.-C. Lee, M. N. Vouvakis, and J.-F. Lee, "A non-overlapping domain decomposition method with non-matching grids for modeling large finite antenna arrays," *J. Comput. Phys.*, vol. 203, no. 1, pp. 1–21, Feb. 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999104003158
[47] J. Lu, Y. Chen, D. Li, and J.-F. Lee, "An embedded domain decomposition method for electromagnetic modeling and design," *IEEE Trans. Antennas Propag.*, vol. 67, no. 1, pp. 309–323, Jan. 2019.

Chung Hyun Lee (Student Member, IEEE) received the B.S. degree in electrical engineering from Inha University, Incheon, South Korea, in 2009, and the M.S. degree in electrical engineering from The Ohio State University, Columbus, OH, USA, in 2011, where he is currently pursuing the Ph.D. degree.

His research interests are numerical techniques in computational electromagnetics.

Joseph D. Kotulski (Life Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Illinois at Chicago, Chicago, IL, USA, in 1975, 1977, and 1983, respectively.

He was at the Naval Research Laboratory, Washington, DC, USA, from 1977 to 1984. Since 1984, he has been a Principal Member of the Technical Staff at Sandia National Laboratories, Albuquerque, NM, USA, where he is currently with the Electromagnetic Theory and Simulation Department. His numerous conference presentations and journal articles encompass a wide range electromagnetic solution techniques applied to a diverse set of problems. His current research interests are in computational electromagnetics using boundary element and finite-element solution techniques on massively parallel platforms as well as accelerated solution algorithms to these equations such as the fast multipole method and other techniques.

Dr. Kotulski is a full member of URSI Commission B.

Vinh Q. Dang (Member, IEEE) received the B.Eng. degree from the Posts and Telecommunications Institute of Technology, Hanoi, Vietnam, in 2003, the M.S. degree from the University of Technology, Vietnam, Ho Chi Minh City, Vietnam, in 2006, and the Ph.D. degree from the Catholic University of America, Washington, DC, USA, in 2015, all in electrical engineering.

Before 2010, he was a Lecturer with the School of Electrical Engineering, International University, Ho Chi Minh City. From 2010 to 2015, he was a Research Assistant with the Electromagnetic Wave Propagation and Remote Sensing Laboratory, Catholic University of America. He was a Research Associate with the Center for Automata Processing, University of Virginia, Charlottesville, VA, USA, from 2015 to 2018. He joined Sandia National Laboratories, Albuquerque, NM, USA, as a Post-Doctoral Appointee in 2018, where he became a Senior Member of Technical Staff in 2019. His research interests include high-performance computing, computational electromagnetics, automata processing, data mining, compressive sensing, radar imaging, and medical image processing.

Jin-Fa Lee (Fellow, IEEE) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1982, and the M.S. and Ph.D. degrees in electrical engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 1986 and 1989, respectively.

From 1988 to 1990, he was with ANSOFT Corporation (later acquired by ANSYS), Pittsburgh, where he was involved in the development of several CAD/CAE finite-element programs for modeling 3-D microwave and millimeter-wave circuits. From 1990 to 1991, he was a Post-Doctoral Fellow with the University of Illinois at Urbana–Champaign, Champaign, IL, USA. From 1991 to 2000, he was with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA, USA. In 2001, he joined The Ohio State University, Columbus, OH, USA, where he is currently a Professor with the Department of Electrical and Computer Engineering. His current research interests include electromagnetic field theories, antennas, numerical methods, and their applications to computational electromagnetics, analyses of numerical methods, fast finite-element methods, fast integral equation methods, hybrid methods, domain decomposition methods, and multiphysics simulations and modeling.