

Diagonal Factorization of Integral Equation Matrices via Localizing Sources and Orthogonally Matched Receivers

Robert J. Adams* and John C. Young

Abstract—A procedure is reported to determine accurate, invertible, block-diagonal factorizations for matrices obtained by discretizing integral equation formulations of electromagnetic interaction problems. The algorithm is based on the combination of localizing source/receiver transformations with orthogonally matched receiver/source transformations. The resulting factorization provides a single, sparse data structure for the system matrix and its inverse, and no approximation is required to convert between the two. Numerical examples illustrate the performance of the factorization for electromagnetic scattering from perfectly conducting elliptical cylinders of different electrical size.

1. INTRODUCTION

The development of efficient simulation tools for electromagnetic radiation, transmission, and scattering problems has been an active area of research for many decades. A relatively recent focus has been the development of sparse, direct solution methods for the linear systems obtained by discretizing time-harmonic integral equation representations of EM interaction problems. Unlike iterative methods, direct methods provide a fixed-cost solution for each excitation. This can provide substantial computational savings relative to iterative methods for problems involving multiple right-hand sides. Sparse direct methods also provide general-purpose preconditioners for fast iterative methods.

Multiple strategies have been used to develop sparse direct solvers, e.g., [1–7]. The set of approaches considered here are based on the use of localizing functions [8–13]. Localization-based direct solvers incorporate source functions having support in individual groups (of an underlying multilevel tree) that localize their scattered fields to the group that contains the source. In one approach, localizing sources have been combined with orthogonally matched receivers to obtain an upper-triangular factorization [10–13]. In the following, this approach is extended to obtain a data-sparse, block-diagonal representation of the system matrix. The new representation is easily inverted without fill-in or further approximation. Numerical examples of TMz scattering from perfect conductors are used to illustrate the performance of the factorization algorithm. A diagonal factorization algorithm for symmetric matrices is detailed in [14].

2. BACKGROUND

Discretized integral equation formulations of linear, time-harmonic, electromagnetic interaction problem can be represented as a matrix equation,

$$\mathbf{Z}\mathbf{x} = \mathbf{F}^i. \quad (1)$$

In this equation, \mathbf{x} is an $N \times k$ matrix containing the coefficients (or degrees of freedom, DOF) of the underlying bases used to represent the sources in the integral equation. The integer N is the number

Received 12 December 2016, Accepted 27 March 2017, Scheduled 13 April 2017

* Corresponding author: Robert J. Adams (robert.adams@uky.edu).

The authors are with the Electrical & Computer Engineering, University of Kentucky, Lexington, KY 40506, USA.

of DOF, and k is the number of excitation vectors. \mathbf{F}^i is an $N \times k$ matrix determined from samples of k excitation functions. The system matrix, \mathbf{Z} , is determined using a moment method to discretize the underlying integral operator [15].

As described in [12, 13], a multilevel tree is used to organize interactions on the simulation domain. The number of levels in the tree is L , and tree levels are indexed from 1 to L . Level-1 is the coarsest (root) level of the tree, and level- L is the finest level. The level index for a general level is level- l , and individual groups at level- l are indexed as $m(l)$. The level index (l) in $m(l)$ will be suppressed for simplicity when the level is clear from the context. In the following, a multilevel, data-sparse, H2 representation [16] is used to represent \mathbf{Z} . The procedure used to construct the H2 representation is discussed in [13].

2.1. Triangular Factorization

Different factorization algorithms can be developed using localizing transformations [8–13]. This section briefly reviews a triangular factorization [10–12]. The proposed diagonal factorization is a modification of this algorithm. In this paper, only non-overlapped localizing functions are used; that is, the support of each localizing function is confined to a single group at a given level of the multilevel tree. (Overlapped localizing functions are necessary for efficient computational performance in 3-D applications [13] and are not considered in this paper.)

The multilevel, triangular factorization obtained using the localizing basis functions is (cf. Equation (27) of [12]),

$$\mathbf{z}_{l+1}^{(NN)} = \mathbf{P}_l \mathbf{P}_l^H \mathbf{z}_{l+1}^{(NN)} \mathbf{\Lambda}_l \mathbf{\Lambda}_l^{-1} \approx \begin{bmatrix} \mathbf{P}_l^{(L)} & \mathbf{P}_l^{(N)} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{z}_l^{(LN)} \\ \mathbf{0} & \mathbf{z}_l^{(NN)} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_l^{(L)} & \mathbf{\Lambda}_l^{(N)} \end{bmatrix}^{-1}, \quad (2)$$

where l indicates the tree level. The original system matrix in Eq. (1) is obtained for $l = L$, $\mathbf{Z} = \mathbf{z}_{L+1}^{(NN)}$. The matrices $\mathbf{\Lambda}_l$ and \mathbf{P}_l in (2) are square, permuted block-diagonal matrices (e.g., see Figure 1(b) of [13]). Their diagonal blocks have a dimension equal to the number of remaining DOF in $\mathbf{z}_{l+1}^{(NN)}$ associated with each level- l group, $m(l)$. The superscripts (L) and (N) are used to indicate matrices associated with localizing and non-localizing sources, respectively. The localizing source DOF are in $\mathbf{\Lambda}_l^{(L)}$, and non-localizing DOF are in $\mathbf{\Lambda}_l^{(N)}$; the latter are selected such that $\mathbf{\Lambda}_l^{(N)}$ is orthogonal to $\mathbf{\Lambda}_l^{(L)}$, which improves the conditioning of $\mathbf{\Lambda}_l$. The matrix \mathbf{P}_l is unitary.

The approximation indicated in Eq. (2) is due to approximate source localization. The localizing source functions in $\mathbf{\Lambda}_l^{(L)}$ only localize their scattered fields, $\mathbf{z}_{l+1}^{(NN)} \mathbf{\Lambda}_l^{(L)}$, to order- ϵ , where ϵ is the desired accuracy for the factorization. It is noted that in (2) and throughout this paper, all matrices $\mathbf{z}_l^{(NN)}$ and $\mathbf{z}_l^{(LN)}$ are represented using an l -level H2 data structure [16]. The procedure used to efficiently manipulate the H2 data structures to compute the localizing functions in $\mathbf{\Lambda}_l^{(L)}$ is discussed in detail in [12, 17]. All other matrices, including $\mathbf{\Lambda}_l$ and \mathbf{P}_l above, and \mathbf{X}_l and \mathbf{Y}_l below, are either block diagonal or permuted block diagonal matrices.

An important property of Eq. (2) is that the submatrix $\mathbf{z}_l^{(NL)} = \mathbf{P}_l^{(N)} \mathbf{z}_{l+1}^{(NN)} \mathbf{\Lambda}_l^{(L)} \approx 0$ is negligible to order- ϵ . This occurs because $\mathbf{P}_l^{(N)}$ is constructed to be orthogonal (to order- ϵ) to the fields radiated by the localizing DOF, $\mathbf{z}_{l+1}^{(NN)} \mathbf{\Lambda}_l^{(L)}$. In this sense, the receiving vectors $\mathbf{P}_l^{(N)}$ are orthogonally matched to the localizing sources, $\mathbf{\Lambda}_l^{(L)}$. This orthogonal matching provides the upper triangular structure of Eq. (2), which enables the following approximate inverse,

$$\mathbf{z}^{-1} = \left(\mathbf{z}_{L+1}^{(NN)} \right)^{-1} \approx \mathbf{\Lambda}_L \begin{bmatrix} \mathbf{I} & -\mathbf{z}_L^{(LN)} \left(\mathbf{z}_L^{(NN)} \right)^{-1} \\ \mathbf{0} & \left(\mathbf{z}_L^{(NN)} \right)^{-1} \end{bmatrix} \mathbf{P}_L^H. \quad (3)$$

The inverse required in Eq. (3) is computed by applying Eqs. (2) and (3) recursively until level-2 is reached. $\mathbf{z}_2^{(NN)}$ is inverted via standard LU decomposition.

3. DIAGONAL FACTORIZATION

The feature of (2) that yields the simple inversion formula (3) is the orthogonal pairing of the receivers in $\mathbf{P}_l^{(N)}$ and the localizing sources in $\mathbf{\Lambda}_l^{(L)}$. In the following, this aspect of (2) is extended to obtain a diagonal factorization for asymmetric matrices. A diagonal factorization algorithm for symmetric matrices is reported in [14].

Begin by defining a right transformation,

$$\mathbf{X}_l = [\mathbf{\Lambda}_{l,R}, \mathbf{P}_{l,R}^*] \quad (4)$$

and a left transformation,

$$\mathbf{Y}_l = [\mathbf{\Lambda}_{l,L}, \mathbf{P}_{l,L}^*] \quad (5)$$

where * denotes complex conjugation. In these equations, the sub-matrix $\mathbf{\Lambda}_{l,R}$ is the sub-matrix $\mathbf{\Lambda}_l^{(L)}$ of Eq. (2), and $\mathbf{P}_{l,L}$ is the sub-matrix $\mathbf{P}_l^{(N)}$. Both $\mathbf{\Lambda}_{l,R}$ and $\mathbf{P}_{l,L}$ are obtained analyzing column blocks of $\mathbf{Z}_l^{(NN)}$ using the algebraic methods detailed in [12, 17]. The sub-matrices $\mathbf{\Lambda}_{l,L}$ and $\mathbf{P}_{l,R}$ are obtained by performing the same algebraic analysis on row-blocks of \mathbf{Z}_l . (It is noted that \mathbf{X}_l and \mathbf{Y}_l are permuted block-diagonal matrices with the same structure as $\mathbf{\Lambda}_l$ and \mathbf{P}_l above.)

The localizing, block-diagonal sub-matrices $\mathbf{\Lambda}_{l,L}$ and $\mathbf{\Lambda}_{l,R}$ can be expressed in terms of their diagonal blocks as

$$\mathbf{\Lambda}_{l,L} = \text{diag}(\boldsymbol{\lambda}_{l,L}^1, \dots, \boldsymbol{\lambda}_{l,L}^m, \dots, \boldsymbol{\lambda}_{l,L}^M) \quad (6)$$

$$\mathbf{\Lambda}_{l,R} = \text{diag}(\boldsymbol{\lambda}_{l,R}^1, \dots, \boldsymbol{\lambda}_{l,R}^m, \dots, \boldsymbol{\lambda}_{l,R}^M) \quad (7)$$

where $\boldsymbol{\lambda}_{l,L}^m$ is the collection of localizing vectors in group- m for which the operation $(\boldsymbol{\lambda}_{l,L}^m)^T \mathbf{Z}_l(m(l), :)$ is nonzero (to order- ϵ) only within columns associated with group- m . (The notation $\mathbf{Z}_l(m(l), :)$ is used to indicate the rows of \mathbf{Z}_l associated with group- m at level- l of the tree.) Similarly, $\boldsymbol{\lambda}_{l,R}^m$ is the collection of localizing vectors in group- m for which the product $\mathbf{Z}_l(:, m(l)) \boldsymbol{\lambda}_{l,R}^m$ is nonzero only within rows contained by the same group- m at level- l of the tree.

The transformations \mathbf{X}_l and \mathbf{Y}_l indicated by Eqs. (4) and (5) provide the following approximation

$$\mathbf{Z}_{l+1} = \mathbf{Y}_l^{-T} \mathbf{Y}_l^T \mathbf{Z}_{l+1} \mathbf{X}_l \mathbf{X}_l^{-1} \approx \mathbf{Y}_l^{-T} \begin{bmatrix} \mathbf{B}_l & 0 \\ 0 & \mathbf{Z}_l \end{bmatrix} \mathbf{X}_l^{-1}, \quad (8)$$

where the system matrix is obtained when $l = L$, $\mathbf{Z} = \mathbf{Z}_{L+1}$. The superscript (NN) in Eq. (2) is dropped for convenience in (8) and in the following. The elements of Eq. (8) are,

$$(\mathbf{P}_{l,L}^*)^T \mathbf{Z}_{l+1} \mathbf{\Lambda}_{l,R} \approx 0, \quad (9)$$

$$(\mathbf{\Lambda}_{l,L})^T \mathbf{Z}_{l+1} \mathbf{P}_{l,R}^* \approx 0, \quad (10)$$

$$(\mathbf{\Lambda}_{l,L})^T \mathbf{Z}_{l+1} \mathbf{\Lambda}_{l,R} = \mathbf{B}_l, \quad (11)$$

$$(\mathbf{P}_{l,L}^*)^T \mathbf{Z}_{l+1} \mathbf{P}_{l,R}^* = \mathbf{Z}_l. \quad (12)$$

The products (9) and (10) are $O(\epsilon)$ for the same reason that $\mathbf{Z}^{(NL)}$ is approximately zero in Eq. (2).

Due to localization, the matrix \mathbf{B}_l is block-diagonal (to order- ϵ):

$$\mathbf{B}_l = \text{diag}(\mathbf{b}_l^1, \dots, \mathbf{b}_l^m, \dots, \mathbf{b}_l^M). \quad (13)$$

The diagonal blocks are

$$\mathbf{b}_l^m = (\boldsymbol{\lambda}_{l,L}^m)^T \mathbf{Z}_l(m(l), m(l)) \boldsymbol{\lambda}_{l,R}^m. \quad (14)$$

Let $n_{l,L}^m$ and $n_{l,R}^m$ denote the number of localizing vectors in $\boldsymbol{\lambda}_{l,L}^m$ and $\boldsymbol{\lambda}_{l,R}^m$. In computing the diagonal blocks of $\mathbf{\Lambda}_{l,L}$ and $\mathbf{\Lambda}_{l,R}$ from the H2 representation of \mathbf{Z}_{l+1} using the algebraic procedure of [12, 17], the number of localizing vectors in $\boldsymbol{\lambda}_{l,L}^m$ and $\boldsymbol{\lambda}_{l,R}^m$ may be different due to the asymmetry of \mathbf{Z}_{l+1} . To maintain square \mathbf{b}_l^m and \mathbf{Z}_l in Eq. (8), the number of localizing vectors retained in $\boldsymbol{\lambda}_{l,L}^m$ and $\boldsymbol{\lambda}_{l,R}^m$ is set equal to the minimum of the initial number of vectors found for each group (i.e., $\min(n_{l,L}^m, n_{l,R}^m)$).

3.1. Error Control

Due to approximation, it is possible that the errors in the factorization of Eq. (8) may be larger than the desired tolerance, $O(\epsilon)$, if the transformations \mathbf{X}_l and/or \mathbf{Y}_l are poorly conditioned. It has been observed that this situation can be avoided by ensuring that the diagonal blocks of the matrix \mathbf{B}_l are sufficiently well-conditioned relative to the tolerance.

An SVD of the diagonal blocks of \mathbf{B}_l provides,

$$\mathbf{b}_l^m = \mathbf{u}_{m(l)} \boldsymbol{\sigma}_{m(l)} \mathbf{v}_{m(l)}^H \quad (15)$$

where $\boldsymbol{\sigma}_{m(l)} = \text{diag}(\sigma_1, \dots, \sigma_k, \dots, \sigma_{K_m})$ is the diagonal matrix of singular values, and K_m indicates the number of localizing sources in group $m(l)$. Let $\bar{\mathbf{v}}_{m(l)}$ ($\bar{\mathbf{u}}_{m(l)}$) denote the columns of $\mathbf{v}_{m(l)}$ ($\mathbf{u}_{m(l)}$) corresponding to singular values that satisfy

$$\sigma_k > \sigma_1 \gamma, \quad (16)$$

where

$$\gamma = \max\left(\epsilon^{1/2}, \kappa_{self}^{-1/2}\right). \quad (17)$$

In Eq. (17), ϵ is the factorization tolerance and κ_{self} is the condition number of the self-block of \mathbf{Z}_{l+1} associated with group $m(l)$ (i.e., the condition number of $\mathbf{Z}_l(m(l), m(l))$). Finally, let $\bar{\boldsymbol{\sigma}}_{m(l)}$ denote the diagonal matrix of singular values associated with $\bar{\mathbf{v}}_{m(l)}$ and $\bar{\mathbf{u}}_{m(l)}$, and let $\bar{\mathbf{s}}_{m(l)}$ denote the diagonal matrix containing the square root of the singular values such that,

$$\bar{\boldsymbol{\sigma}}_{m(l)} = \bar{\mathbf{s}}_{m(l)} \bar{\mathbf{s}}_{m(l)} \quad (18)$$

The block-diagonal collection of the $\bar{\mathbf{v}}_{m(l)}$ ($\bar{\mathbf{u}}_{m(l)}$) for all level- l groups is denoted $\bar{\mathbf{v}}_l$ ($\bar{\mathbf{u}}_l$). Similarly, the collection of the $\bar{\mathbf{s}}_{m(l)}$ is denoted $\bar{\mathbf{s}}_l$. Updated versions of Eqs. (4) and (5) that provide better conditioning are thus

$$\bar{\mathbf{X}}_l = [\bar{\boldsymbol{\Lambda}}_{l,R}, \bar{\mathbf{P}}_{l,R}^*] \quad (19)$$

$$\bar{\mathbf{Y}}_l = [\bar{\boldsymbol{\Lambda}}_{l,L}, \bar{\mathbf{P}}_{l,L}^*] \quad (20)$$

where $\bar{\boldsymbol{\Lambda}}_{l,R} = \boldsymbol{\Lambda}_{l,R} \bar{\mathbf{v}}_l \bar{\mathbf{s}}_l^{-1}$ and $\bar{\boldsymbol{\Lambda}}_{l,L} = \boldsymbol{\Lambda}_{l,L} \bar{\mathbf{u}}_l^* \bar{\mathbf{s}}_l^{-1}$. The diagonal blocks of $\bar{\mathbf{P}}_{l,R}$ (or $\bar{\mathbf{P}}_{l,L}$) are determined from the null-space of the localized fields, $\mathbf{Z}_{l+1} \bar{\boldsymbol{\Lambda}}_{l,R}$ (or $\bar{\boldsymbol{\Lambda}}_{l,L}^T \mathbf{Z}_{l+1}$), in the manner described for $\mathbf{P}_l^{(N)}$ of Eq. (2) in [10, 12].

3.2. Multilevel Factorization

The updated version of Eq. (8) obtained using Eqs. (19) and (20) is,

$$\mathbf{Z}_{l+1} \approx \bar{\mathbf{Y}}_l^{-T} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{Z}_l \end{bmatrix} \bar{\mathbf{X}}_l^{-1}. \quad (21)$$

As shown in the following numerical examples, the error-controlled factorization in Eq. (21) is significantly more accurate than the factorization in Eq. (8). The multilevel approximate inverse associated with Eq. (21) is,

$$\mathbf{Z}_{l+1}^{-1} \approx \bar{\mathbf{X}}_l \begin{bmatrix} I & 0 \\ 0 & \mathbf{Z}_l^{-1} \end{bmatrix} \bar{\mathbf{Y}}_l^T. \quad (22)$$

The recursive expansion of the factorization in Eq. (21) is,

$$\mathbf{Z} \approx \bar{\mathbf{Y}}_L^{-T} \bar{\mathbf{Y}}_{L-1}^{-T} \dots \bar{\mathbf{Y}}_3^{-T} \bar{\mathbf{Y}}_2^{-T} \begin{bmatrix} I & 0 \\ 0 & \mathbf{Z}_2 \end{bmatrix} \bar{\mathbf{X}}_2^{-1} \bar{\mathbf{X}}_3^{-1} \dots \bar{\mathbf{X}}_{L-1}^{-1} \bar{\mathbf{X}}_L^{-1}, \quad (23)$$

and the recursive expansion of Eq. (22) is

$$\mathbf{Z}^{-1} \approx \bar{\mathbf{X}}_L \bar{\mathbf{X}}_{L-1} \dots \bar{\mathbf{X}}_3 \bar{\mathbf{X}}_2 \begin{bmatrix} I & 0 \\ 0 & \mathbf{Z}_2^{-1} \end{bmatrix} \bar{\mathbf{Y}}_2^T \bar{\mathbf{Y}}_3^T \dots \bar{\mathbf{Y}}_{L-1}^T \bar{\mathbf{Y}}_L^T. \quad (24)$$

From Eqs. (23) and (24) it is observed that, in addition to providing a block-diagonal inverse approximation, the diagonal factorization in Eq. (21) also provides a block-diagonal representation of \mathbf{Z} (each $\bar{\mathbf{X}}_l$ and $\bar{\mathbf{Y}}_l$ is a block-diagonal matrix at level- l , and \mathbf{Z}_2 is a square matrix with dimension equal to the number of remaining, non-localized DOF). Note that \mathbf{Z} and \mathbf{Z}^{-1} have identical block-diagonal structures. Thus, once the diagonal representation in Eq. (23) is determined, the inverse Eq. (24) can be obtained without additional approximation or fill-in. In this sense, Eq. (23) provides a single, explicit, data-sparse representation of both the system matrix, \mathbf{Z} , and its inverse, \mathbf{Z}^{-1} . Although the triangular factorization in Eq. (2) can be inverted via Eq. (3), the inverse requires back-substitution and application of $\mathbf{Z}_l^{(LN)}$ which is stored as an l -level H2 data structure, at each level of the solution process. For this reason, when using Eq. (3), the memory required by the factorization exceeds the memory used by the initial H2 representation of \mathbf{Z} , and the solve process is slowed by the need to apply an l -level H2 data structure at each level of the solve when applying $\mathbf{Z}_l^{(LN)}$.

In contrast, using the inverse Eq. (24), it is not necessary to store any off-diagonal data or perform any back-substitution during the solution process. As shown below, the memory required by the diagonal representation in Eq. (24) is actually less than that used by the initial H2 representation in some cases.

3.3. Integrated Factor and Solve (IFS)

An advantage of the diagonal representation in Eq. (24) is that, if the excitation \mathbf{F}^i is known *a priori*, an integrated factor and solve (IFS) can be used to reduce the overall memory requirement of the sparse direct solver.

The IFS procedure proceeds as follows. At each level of the factorization, $\bar{\mathbf{Y}}_l$ can be applied to the right-hand side vectors, \mathbf{F}^i , as indicated by Eq. (24) and then deallocated. Similarly, at each level the localizing submatrix of $\bar{\mathbf{X}}_l$, $\bar{\Lambda}_{l,R}$, can be applied and then deallocated. The blocks $\bar{\mathbf{P}}_{l,R}^*$ are the only blocks of the diagonal factorization retained throughout the entire IFS process. Once the factorization is completed, the block \mathbf{Z}_2 is factored and applied to the right-hand side vector(s), after which the $\bar{\mathbf{P}}_{l,R}^*$ are finally applied, proceeding from coarse to fine levels of the tree, to determine the solution vectors, \mathbf{x} .

4. NUMERICAL RESULTS

To illustrate the numerical properties of the diagonal factorization, consider the problem of TMz scattering from perfectly conducting elliptical cylinders. The magnetic field integral equation (MFIE) is used to formulate the problem. The matrix equation (1) is obtained from the MFIE using a point-matching moment method with pulse basis functions [15] for the case of scattering from a pair of ellipses separated by 0.1λ , each with an axial ratio of 4. The major axis, D , of the ellipses varies between 12λ and 1536λ . The scatterer configuration for $D = 1536\lambda$ is shown in Figure 1. In all cases considered, more than ten basis functions are used per linear wavelength. Similar numerical experiments are reported for symmetric and asymmetric matrices in [14].

The factorization algorithm outlined above relies on the determination of the localizing DOF used to define the blocks $\Lambda_{l,R}$ and $\Lambda_{l,L}$. As discussed elsewhere [12,13], this is accomplished by first building a data-sparse H2 representation of \mathbf{Z} [16]. The methods used to efficiently manipulate the H2 representation to determine the submatrices $\Lambda_{l,L}$, $\Lambda_{l,R}$, $\mathbf{P}_{l,L}$, and $\mathbf{P}_{l,R}$ of $\bar{\mathbf{X}}_l$ and $\bar{\mathbf{Y}}_l$ are detailed in [12,17]. When the error control procedure is used, the modification indicated by Eq. (15) through Eq. (20) is used to replace \mathbf{X}_l and \mathbf{Y}_l by $\bar{\mathbf{X}}_l$ and $\bar{\mathbf{Y}}_l$. In the following examples, the initial H2 representation of \mathbf{Z} is constructed using a tolerance of 1E-6; the subsequent diagonal factorization is performed using a tolerance of 1E-5.

Tables 1 and 2 report relative RMS matrix errors in the H2 data structure and the diagonal factorization. Table 1 reports errors for a fixed $D = 12\lambda$; N is increased by increasing the density of the discretization per wavelength. Table 2 reports the matrix errors for the cases of a fixed discretization density; N is increased by increasing the electrical sizes of the ellipses.

In all cases in the tables, the H2 representation of \mathbf{Z} is near the specified tolerance of 1E-6. The error in the factorization is similarly well controlled when the proposed error control procedure is used; the error is less than the factorization tolerance of 1E-5 in all cases. It is also observed that much larger

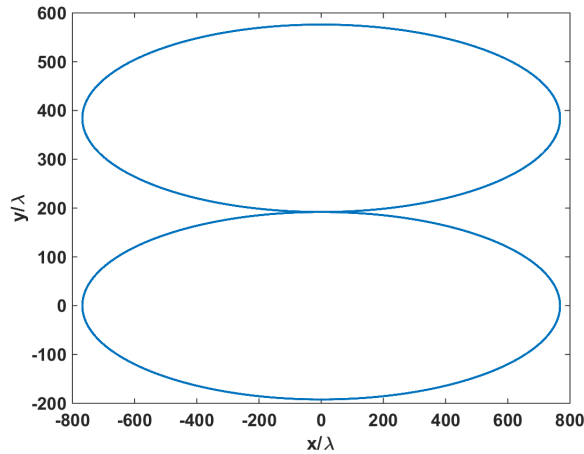


Figure 1. Two-ellipse scattering configuration when $D = 1536\lambda$ ($N = 135112$). The ellipses are separated by 0.1λ at their point of closest proximity.

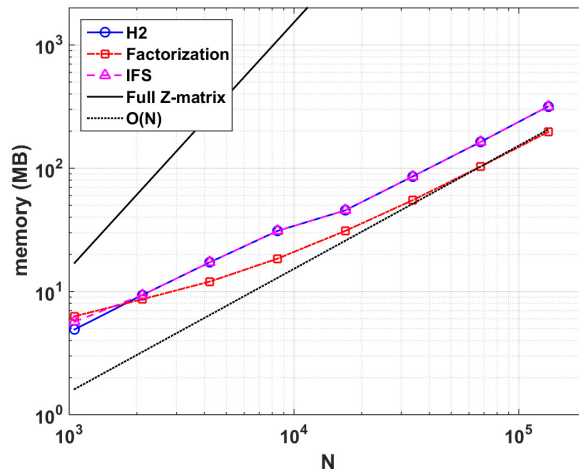


Figure 2. Memory used by H2 representation, diagonal factorization and integrated factor/solve (IFS) procedure for fixed $D = 12\lambda$ and increasing discretization density. Also shown are $O(N)$ and $O(N^2)$ reference lines.

factorization errors occur when the error control strategy is not employed. For this reason, the error control method outlined above is used in all remaining examples.

Figure 2 displays the memory used for the H2 representation, for the factorization, and for the integrated factor/solve (IFS) procedure for the same fixed major axis cases considered in Table 1. It is observed that the memory requirements increase approximately linearly. Furthermore, the diagonal factorization requires less memory than the H2 representation for all but the smallest case ($N = 1056$). This suggests that for such cases it may be advantageous to develop a method to directly construct the diagonal representation in Eq. (23) rather than first building the H2 structure as was done here. Finally, it is observed that the peak memory required by the IFS procedure does not notably exceed the peak memory required by the initial H2 data structure for these cases.

Figure 3 reports the memory requirements when the discretization density is fixed and the major axis, D , increases (cf. Table 2). In these cases, the memory requirements all increase approximately as $O(N^{1.5})$, which is substantially less than the $O(N^2)$ cost to store the full matrix, \mathbf{Z} . Unlike the cases reported in Figure 2, the memory required by the diagonal factorization exceeds that used by the H2 representation. However, the IFS procedure limits the amount by which the factorization procedure causes the peak memory to exceed that required to store the initial H2 representation. Using IFS, the

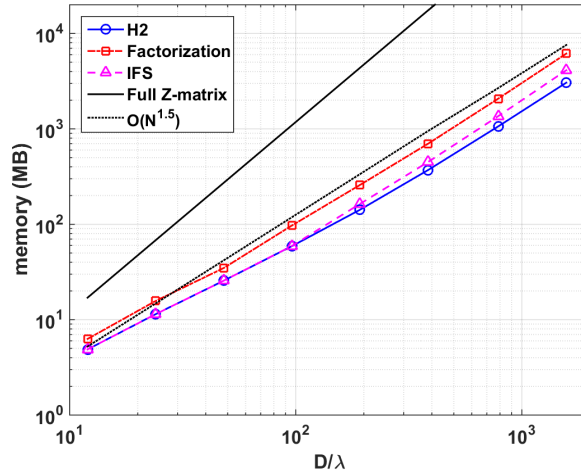


Figure 3. Memory used by H2 representation, diagonal factorization, and IFS for fixed discretization density and increasing major axis length. Also shown are $O(N^{1.5})$ and $O(N^2)$ reference lines.

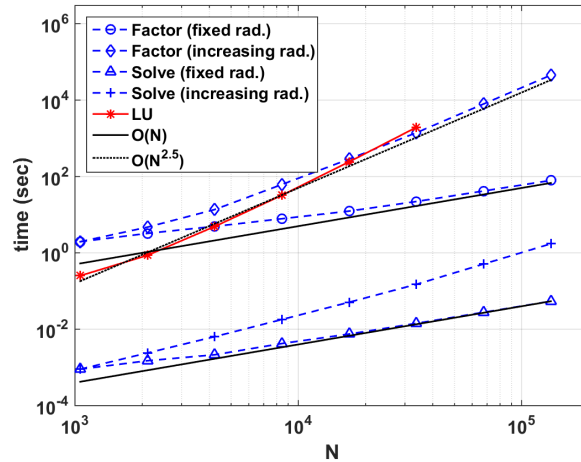


Figure 4. Diagonal factorization and solve times for cases of increasing discretization density with fixed $D = 12\lambda$, and for cases of fixed discretization density with increasing D . The performance of an optimized LU factorization is also shown.

increase caused by the direct solver only becomes notable for values of D larger than about 100λ .

Finally, Figure 4 indicates the time used to perform the factorization for the cases in Tables 1 and 2. (All computations were performed on an Intel Xeon X5450 3 GHz Quad-Core Processor.) The figure also shows the time required to solve for a single right-hand side vector given the completed diagonal factorization. In the case of a fixed electrical size and increasing discretization density, the factorization and solution times scale approximately linearly, and the diagonal factorization substantially outperforms LU factorization. However, in the case of a fixed discretization density with increasing D , the factorization does not begin to outperform LU factorization until N exceeds about 16000.

4.1. Other Formulations

Electromagnetic scattering problems can be formulated using a variety of integral equations in addition to the MFIE used here. It has been observed that the CFIE formulation [15] yields numerical results that are essentially similar to those reported here for the MFIE. In the case of an EFIE, the factorization errors (cf. Tables 1 and 2) are also similar. However, larger computational complexity has been observed

Table 1. Relative RMS errors in \mathbf{Z} of (1) using H2 representation and the diagonal factorization (23) with and without error control (EC) procedure. The electrical size is fixed at $D = 12\lambda$ and N is increased by increasing the discretization density. Values > 1 indicate that the error exceeds 100 percent. The number of quad-tree levels varies from $L = 4$ ($N = 1056$) to $L = 11$ ($N = 135112$).

D/λ (N)	H2	Factor with EC	Factor w/o EC
12 (1056)	3.6E-7	1.8E-6	> 1
12 (2112)	9.8E-6	9.8E-6	> 1
12 (4222)	3.0E-7	1.6E-6	> 1
12 (8444)	2.9E-7	1.4E-6	> 1
12 (16890)	2.6E-7	1.3E-6	0.956
12 (33778)	2.5E-7	1.5E-6	0.664
12 (67556)	2.9E-7	1.2E-6	0.477
12 (135112)	2.8E-7	1.3E-6	0.412

Table 2. Same as Table 1 but with increasing electrical size (D) and fixed discretization density.

D/λ (N)	H2	Factor with EC	Factor w/o EC
12 (1056)	3.6E-7	1.8E-6	5.2E-4
24 (2112)	3.4E-6	4.0E-6	> 1
48 (4222)	5.3E-7	1.7E-6	> 1
96 (8444)	3.6E-6	3.9E-6	> 1
192 (16890)	8.3E-7	1.6E-6	> 1
384 (33778)	4.2E-7	1.4E-6	> 1
768 (67556)	1.8E-6	2.2E-6	> 1
1536 (135112)	1.1E-6	1.9E-6	> 1

when the EFIE is used to solve the problems reported herein. This was previously observed in [13] for the EFIE and is due to the well-known ill-conditioning of the EFIE. Fortunately, better conditioned formulations are available, including the augmented EFIE [18, 19]. It has previously been shown that the computational complexity of the factorization (2) is $O(N \log N)$ or better for an augmented EFIE formulation in three dimensions [20]. The application of the new, diagonal factorization to such formulations will be a subject of future investigation.

5. SUMMARY

A diagonal, data-sparse factorization strategy for asymmetric integral equation matrices has been detailed, and a computationally efficient error control strategy has been proposed. The factorization provides a single, sparse data structure for both the system matrix (23) and its inverse (24). There is no fill-in or additional approximation required to convert between the diagonal representations of \mathbf{Z} and \mathbf{Z}^{-1} . Due to the diagonal nature of the representation, the solution procedure used to determine \mathbf{x} from \mathbf{F}^i involves only explicit matrix multiplication.

The numerical properties of the algorithm have been illustrated for an MFIE formulation of electromagnetic scattering from perfectly conducting elliptical cylinders. In cases for which the electrical size is fixed and N is increased by increasing the discretization density, the computational costs of the proposed algorithm scale approximately linearly. In such cases, the diagonal factorization uses less memory than the original H2 representation of \mathbf{Z} . This suggests that it may be beneficial to fill the diagonal representation directly in some cases. When the discretization density is fixed and the

electrical size is increased, the diagonal factorization generally uses more memory than the initial H2 representation of \mathbf{Z} . However, the additional memory required by the diagonal factorization can be reduced using an integrated factor/solve (IFS) procedure.

For three-dimensional surface integral equation applications with non-oscillatory kernels, the computational complexity of the algorithm reported here is limited to $O(N^{1.5})$ due to the use of non-overlapped localizing functions [11, 12]. However, the factorization algorithm admits a straightforward incorporation of the overlapped functions defined in [13]. These functions have been observed to yield factorization complexities that scale as $O(N \log N)$ or better for surface integral equations in 3-D using the triangular factorization of (2) [13, 20]. The incorporation of overlapped localizing functions within the diagonal factorization outlined here will be a subject of future investigation.

ACKNOWLEDGMENT

The authors gratefully acknowledge support from the Office of Naval Research (N00014-16-1-3066) and the Kentucky Science and Engineering Foundation (KSEF-148-502-16-393).

REFERENCES

1. Gope, D., I. Chowdhury, and V. Jandhyala, "DiMES: Multilevel fast direct solver based on multipole expansions for parasitic extraction of massively coupled 3D microelectronic structures," *42nd Design Automation Conference*, 159–162, 2005.
2. Shaeffer, J., "Direct solve of electrically large integral equations for problems sizes to 1 M unknowns," *IEEE Transactions on Antennas and Propagation*, Vol. 56, 2306–2313, 2008.
3. Jiao, D. and W. Chai, "An H2-matrix-based integral-equation solver of reduced complexity and controlled accuracy for solving electrodynamic problems," *IEEE Transactions on Antennas and Propagation*, Vol. 57, 3147–3159, 2009.
4. Heldring, A., J. M. Rius, J. M. Tamayo, J. Parron, and E. Ubeda, "Multiscale compressed block decomposition for fast direct solution of method of moments linear system," *IEEE Transactions on Antennas and Propagation*, Vol. 59, 526–536, 2011.
5. Wei, J.-G., 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 Science*, Vol. 47, 1–9, 2012.
6. Brick, Y., V. Lomakin, and A. Boag, "Fast direct solver for essentially convex scatterers using multilevel non-uniform grids," *IEEE Transactions on Antennas and Propagation*, Vol. 62, 4314–4324, 2015.
7. Zhang, Y. L., X. Chen, C. Fei, Z. Li, and C. Gu, "Fast direct solution of composite conducting-dielectric arrays using Sherman-Morrison-Woodbury algorithm," *Progress In Electromagnetics Research M*, Vol. 49, 203–209, 2016.
8. Martinsson, P. G. and V. Rokhlin, "A fast direct solver for boundary integral equations in two dimensions," *Journal of Computational Physics*, Vol. 205, 1–23, 2005.
9. Adams, R. J., A. Zhu, and F. X. Canning, "Efficient solution of integral equations in a localizing basis," *Journal of Electromagnetic Waves and Applications*, Vol. 19, No. 12, 1583–1594, 2005.
10. Zhu, A., R. J. Adams, F. X. Canning, and S. D. Gedney, "Schur factorization of the impedance matrix in a localizing basis," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 3, 351–362, 2006.
11. Adams, R. J., A. Zhu, and F. X. Canning, "Sparse factorization of the TMz impedance matrix in an overlapped localizing basis," *Progress In Electromagnetics Research*, Vol. 61, 291–322, 2006.
12. Adams, R. J., Y. Xu, X. Xu, J. S. Choi, S. D. Gedney, and F. X. Canning, "Modular fast direct electromagnetic analysis using local-global solution modes," *IEEE Transactions on Antennas and Propagation*, Vol. 56, 2427–2441, Aug. 2008.
13. Xu, X. and R. J. Adams, "Sparse matrix factorization using overlapped localizing LOGOS modes on a shifted grid," *IEEE Transactions on Antennas and Propagation*, Vol. 60, 1414–1424, 2012.

14. Adams, R. J. and J. C. Young, "A diagonal factorization for integral equation matrices," *2016 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 812–815, September 19–23, 2016.
15. Peterson, A. F., S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*, IEEE Press, Piscataway, NJ, 1998.
16. Hackbusch, W. and S. Borm, "H2-matrix approximation of integral operators by interpolation," *Applied Numerical Mathematics*, Vol. 43, 129–143, 2002.
17. Xu, Y., X. Xu, and R. J. Adams, "A sparse factorization for fast computation of localizing modes," *IEEE Transactions on Antennas and Propagation*, Vol. 58, 3044–3049, 2010.
18. Qian, Z. G. and W. C. Chew, "An augmented electric field integral equation for highspeed interconnect analysis," *Microwave and Optical Technology Letters*, Vol. 50, 2658–2662, October 2008.
19. Qian, Z. G. and W. C. Chew, "Enhanced A-EFIE with perturbation method," *IEEE Transactions on Antennas and Propagation*, Vol. 58, 3256–3264, October 2010.
20. Cheng, J. and R. J. Adams, "Direct solution method using overlapped localizing LOGOS modes for AEFIE-G at low frequencies," *28th Annual Review of Progress in Applied Computational Electromagnetics*, 579–584, Columbus, Ohio, 2012.