

Preconditioning Techniques for Domain Integral Equations

R.F. Remis*

Abstract — In this paper, we consider so-called optimal circulant preconditioners for discretized integral operators describing the scattering of steady state electromagnetic waves by penetrable objects embedded in homogeneous background media. For two-dimensional scattering problems, we show that the preconditioners may significantly increase the convergence rate of an iterative solver. Possible extensions are discussed as well and two numerical examples illustrate the performance of the preconditioner.

1 INTRODUCTION

We consider scattering of steady-state electromagnetic waves by penetrable bounded objects embedded in a homogeneous background medium. For one- and two-dimensional configurations, we formulate this scattering problem in terms of scalar integral equations. It is well known, of course, that the integral operators have a convolution structure in these cases, since the background medium is linear and shift-invariant. As a consequence, the matrices that result after a standard discretization procedure on a uniform grid have a (recursive) Toeplitz structure. This structure allows us to compute matrix-vector products via FFT and this is one of the reasons why such systems are often solved using an iterative solver.

Efficient preconditioners may significantly speed up the convergence of iterative solvers and circulant preconditioners are particularly effective for a specific class of Toeplitz systems [1]. A popular choice is the so-called optimal circulant preconditioner as introduced in [2]. This preconditioner is uniquely defined as the best circulant fit to a given (Toeplitz) matrix as measured in the Frobenius norm. The preconditioner clusters the eigenvalues around one and its action on a vector can be computed via FFT. Furthermore, its elements can be given explicitly in terms of the elements of the Toeplitz matrix.

In [3], we showed that the optimal preconditioner is very effective for one-dimensional homogeneous objects having Toeplitz matrix representations and significantly speeds up the convergence of an iterative solver. Moreover, for inhomogeneous objects

we showed that explicit optimal circulant preconditioners can be determined as well, even though Toeplitz structure is lost in this case.

In this paper, we extend the approach followed for one-dimensional problems to two-dimensional scalar wave field problems. Specifically, we show that for homogeneous objects represented by matrices with block Toeplitz structure, the optimal block preconditioner again significantly reduces the number of iterations and computation times. Moreover, we show that the analysis for inhomogeneous objects (for which Toeplitz structure is lost) can be extended to the two-dimensional case as well. Finally, two numerical examples will illustrate the performance of the block optimal preconditioner.

2 DOMAIN INTEGRAL EQUATIONS IN ONE DIMENSION

In this section, we briefly review the construction of the optimal circulant preconditioner for the one-dimensional case, since it serves as a building block for two-dimensional problems.

We consider a one-dimensional configuration showing variation in the y -direction only. A penetrable slab with medium parameters $\sigma^{\text{sc}}(y)$ and $\varepsilon^{\text{sc}}(y)$ occupies the domain $\mathbb{D}^{\text{sc}} = \{y \in \mathbb{R}; 0 < y < d\}$. The slab is embedded in a homogeneous background medium with constant medium parameters σ and ε and we set $\eta = \sigma + j\omega\varepsilon$ and $\zeta = j\omega\mu$. In this 1D configuration, the electric field strength E_z satisfies the integral equation

$$E_z(y) - k_b^2 \int_{y'=0}^d G(y-y')\chi(y')E_z(y') dy' = E_z^{\text{inc}}(y), \quad (1)$$

with $y \in \mathbb{D}^{\text{sc}}$. Here, E_z^{inc} is the known incident electric field strength and k_b is the wave number of the background medium defined as $k_b = (-\eta\zeta)^{1/2}$ with $\text{Im}(k_b) \leq 0$. Furthermore, χ is the contrast function given by $\chi(y) = \eta^{\text{sc}}(y)/\eta - 1$, where $\eta^{\text{sc}}(y) = \sigma^{\text{sc}}(y) + j\omega\varepsilon^{\text{sc}}(y)$, and G is the Greens function of the homogeneous background medium.

Using the composite midpoint rule to discretize the integral and applying a standard point matching procedure, we arrive at the discretized system $\mathbf{K}\mathbf{e}_z = \mathbf{e}_z^{\text{inc}}$ with $\mathbf{K} = \mathbf{I} - \mathbf{G}\mathbf{X}$. In this equation, the unknown total field and known incident

*Circuits and Systems Group, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands, e-mail: R.F.Remis@tudelft.nl.

field values at the $N \geq 1$ grid nodes y_n are stored in the vectors $\mathbf{e}_z = [E_z(y_1), E_z(y_2), \dots, E_z(y_N)]^T$ and $\mathbf{e}_z^{\text{inc}} = [E_z^{\text{inc}}(y_1), E_z^{\text{inc}}(y_2), \dots, E_z^{\text{inc}}(y_N)]^T$, respectively, while the diagonal contrast matrix is given by $\mathbf{X} = \text{diag}(\chi_1, \chi_2, \dots, \chi_N)$. Finally, the elements on the n th diagonal of the Green's matrix \mathbf{G} are given by $g_n = k_b \delta y / (2j) \exp(-jk_b \delta y |n|)$ for $n = 0, \pm 1, \pm 2, \dots, \pm(N-1)$ and $\delta y > 0$ is the step size. Notice that \mathbf{G} is a complex symmetric Toeplitz matrix of order N and its action on a vector can be computed via FFTs.

Now instead of solving $\mathbf{K}\mathbf{e}_z = \mathbf{e}_z^{\text{inc}}$ directly, we solve the preconditioned system $\mathbf{M}^{-1}\mathbf{K}\mathbf{e}_z = \mathbf{M}^{-1}\mathbf{e}_z^{\text{inc}}$. Here, \mathbf{M} is the preconditioner and we take it to be of the form $\mathbf{M} = \mathbf{I} - \mathbf{C}$, where \mathbf{C} is a circulant matrix of order N . Specifically, we follow [2] and take

$$\mathbf{C} = \underset{\mathbf{Z} \text{ circulant}}{\text{argmin}} \|\mathbf{G}\mathbf{X} - \mathbf{Z}\|_{\text{F}}, \quad (2)$$

where $\|\cdot\|_{\text{F}}$ denotes the Frobenius norm. This circulant is known as the optimal circulant preconditioner [2]. Writing the first column of matrix \mathbf{C} as $\mathbf{c} = [c_0, c_1, \dots, c_{N-1}]^T$ and introducing the commulative contrast values

$$s_i = \sum_{j=1}^{N-i} \chi_j \quad \text{for } i = 0, 1, \dots, N-1, \quad (3)$$

the elements of the optimal circulant are given by [3]

$$c_i = \frac{g_{N-i}(s_0 - s_i) + g_i s_i}{N} \quad \text{for } i = 0, 1, \dots, N-1. \quad (4)$$

Having found the elements of the circulant, we exploit the property that a circulant matrix is diagonalized by the unitary DFT matrix \mathbf{F} . In particular, we have $\mathbf{C} = \mathbf{F}^H \mathbf{D} \mathbf{F}$, where $\mathbf{D} = \sqrt{N} \text{diag}(\mathbf{F}\mathbf{c})$ is a diagonal matrix with the eigenvalues of matrix \mathbf{C} on its diagonal. With the help of this eigendecomposition, we can write $\mathbf{M}^{-1} = \mathbf{F}^H (\mathbf{I} - \mathbf{D})^{-1} \mathbf{F}$, showing that the action of \mathbf{M}^{-1} on a vector can be computed at FFT speed as well. Finally, we mention that the optimal circulant can also be written as $\mathbf{C} = \mathbf{F}^H \mathbf{d} (\mathbf{F} \mathbf{G} \mathbf{X} \mathbf{F}^H) \mathbf{F}$, where $\mathbf{d}(\mathbf{A})$ is a diagonal matrix with the diagonal entries of matrix \mathbf{A} on its diagonal [2, 3].

3 SCALAR DOMAIN INTEGRAL EQUATIONS IN TWO DIMENSIONS

Having reviewed the one-dimensional case, we now consider E-polarized fields in a configuration that is invariant in the z -direction. An object occupies a bounded domain \mathbb{D}^{sc} in the transverse plane

and is characterized by the medium parameters $\eta^{\text{sc}}(x, y) = \sigma^{\text{sc}}(x, y) + j\omega\epsilon^{\text{sc}}(x, y)$ and $\zeta = j\omega\mu$. The object is embedded in a homogeneous background medium with medium parameters $\eta = \sigma + j\omega\epsilon$ and $\zeta = j\omega\mu$ and the wave number of the background medium is again denoted by k_b .

It is well known that the electric field strength E_z inside the scattering domain satisfies the integral equation

$$E_z(\mathbf{x}) - A_z(\mathbf{x}) = E_z^{\text{inc}}(\mathbf{x}) \quad \text{with } \mathbf{x} \in \mathbb{D}^{\text{sc}}. \quad (5)$$

In the above equation, the vector potential A_z is given by

$$A_z(\mathbf{x}) = k_b^2 \int_{\mathbf{x}' \in \mathbb{D}^{\text{sc}}} G(\mathbf{x} - \mathbf{x}') \chi(\mathbf{x}') E_z(\mathbf{x}') dA, \quad (6)$$

and G is the two-dimensional Greens function. Note that we have included the factor k_b^2 in our definition of the vector potential.

To discretize the above integral equation, we take a rectangular scattering domain with side lengths ℓ_x and ℓ_y in the x - and y -direction, respectively. The upper left corner of this domain coincides with the origin of our Cartesian reference frame. Furthermore, we introduce the grid node coordinates $x_m = \frac{\delta x}{2} + (m-1)\delta x$ for $m = 1, 2, \dots, M$ and $y_n = \frac{\delta y}{2} + (n-1)\delta y$ for $n = 1, 2, \dots, N$, where $\delta x = \ell_x/M$ and $\delta y = \ell_y/N$. The position vector of a grid node is given by $\mathbf{x}_{mn} = x_m \mathbf{i}_x + y_n \mathbf{i}_y$, for $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$.

The scattering domain is divided into nonoverlapping discretization cells $\mathbb{S}_{ij} = \{(i-1)\delta x < x < i\delta x, (j-1)\delta y < y < j\delta y\}$ for $i = 1, 2, \dots, M$ and $j = 1, 2, \dots, N$. Finally, we assume that the contrast function is constant in each cell, that is, $\chi(x_m, y_n) = \chi_{mn}$ for $m = 1, 2, \dots, M$ and $j = 1, 2, \dots, N$, where the χ_{mn} are constants.

As a first step of the discretization procedure, we require that Eq. (5) holds at the grid nodes. We have $E_z(x_m, y_n) - A_z(x_m, y_n) = E_z^{\text{inc}}(x_m, y_n)$ for $m = 1, 2, \dots, M$ and $n = 1, 2, \dots, N$. Introducing the matrices \mathbf{E}_z , \mathbf{A}_z , and $\mathbf{E}_z^{\text{inc}}$ with elements $\mathbf{E}_z|_{mn} = E_z(x_m, y_n)$, $\mathbf{A}_z|_{mn} = A_z(x_m, y_n)$, and $\mathbf{E}_z^{\text{inc}}|_{mn} = E_z^{\text{inc}}(x_m, y_n)$ for $m = 1, 2, \dots, M$ and $n = 1, 2, \dots, N$, we can write the above equations more compactly as $\mathbf{E}_z - \mathbf{A}_z = \mathbf{E}_z^{\text{inc}}$. Applying the vec-operation to this equation and using the linearity of this operator, we arrive at

$$\text{vec}(\mathbf{E}_z) - \text{vec}(\mathbf{A}_z) = \text{vec}(\mathbf{E}_z^{\text{inc}}). \quad (7)$$

As a final step, we need to relate the vector potential $\text{vec}(\mathbf{A}_z)$ to the electric field strength $\text{vec}(\mathbf{E}_z)$. The vector potential is related

to the electric field strength by Eq. (6). This integral cannot be discretized in a straightforward manner, because the Hankel function has a logarithmic singularity at the origin. To remedy this problem, we compute the vector potential using a weakened Greens function denoted by G^w . More precisely, the vector potential is computed using

$$A_z(\mathbf{x}) = k_b^2 \int_{\mathbf{x}' \in \mathbb{D}^{\text{sc}}} G^w(\mathbf{x} - \mathbf{x}') \chi(\mathbf{x}') E_z(\mathbf{x}') dA, \quad (8)$$

where the weakened Greens function is given by

$$G^w(\mathbf{x}) = \frac{1}{\pi a^2} \int_{\mathbf{x}' \in \mathbb{D}^{\text{circ}}} G(\mathbf{x} - \mathbf{x}') dA, \quad (9)$$

and \mathbb{D}^{circ} is a circular disk of radius $a = \frac{1}{2} \min\{\delta x, \delta y\}$.

Now using the fact that the contrast function is constant in each discretization cell and discretizing the integral using the composite midpoint rule, we obtain

$$A_z(x_m, y_n) \approx (k_b \delta x)(k_b \delta y) \sum_{i=1}^M \sum_{j=1}^N G^w(\mathbf{x}_{mn} - \mathbf{x}_{ij}) \chi_{ij} E_z(x_i, y_j), \quad (10)$$

for $m = 1, 2, \dots, M$ and $n = 1, 2, \dots, N$.

We are now in a position to write Eq. (10) in matrix-vector notation. Specifically, introducing the diagonal contrast matrix given by (using Matlab notation)

$$\mathbf{X} = \text{diag}(\chi_{:,1}) \oplus \text{diag}(\chi_{:,2}) \oplus \dots \oplus \text{diag}(\chi_{:,N}), \quad (11)$$

we can write Eq. (10) as

$$\text{vec}(\mathbf{A}_z) = \mathbf{G}\mathbf{X}\text{vec}(\mathbf{E}_z), \quad (12)$$

where the spatial convolution matrix \mathbf{G} has a block matrix structure given by

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}^{(1,1)} & \mathbf{G}^{(1,2)} & \dots & \mathbf{G}^{(1,N)} \\ \mathbf{G}^{(2,1)} & \mathbf{G}^{(2,2)} & \dots & \mathbf{G}^{(2,N)} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{G}^{(N,1)} & \mathbf{G}^{(N,2)} & \dots & \mathbf{G}^{(N,N)} \end{pmatrix} \quad (13)$$

and the (m, i) element of block matrix $\mathbf{G}^{(n,j)}$ is given by

$$\mathbf{G}^{(n,j)} \Big|_{mi} = (k_b \delta x)(k_b \delta y) G^w(\mathbf{x}_{mn} - \mathbf{x}_{ij}). \quad (14)$$

From the above expressions, it is easily verified that the spatial convolution matrix \mathbf{G} is complex-symmetric, block Toeplitz, and each block matrix $\mathbf{G}^{(n,j)}$ is again Toeplitz (matrix \mathbf{G} is a so-called BTTB matrix).

Substitution of Eq. (12) in Eq. (7), we arrive at the system $\mathbf{K}\text{vec}(\mathbf{E}_z) = \text{vec}(\mathbf{E}_z^{\text{inc}})$, where $\mathbf{K} = \mathbf{I} - \mathbf{G}\mathbf{X}$ and \mathbf{I} is the identity matrix of order MN .

Before we discuss how we construct the preconditioner, we first introduce a permutation matrix that relates vector $\text{vec}(\mathbf{E}_z)$ to vector $\text{vec}(\tilde{\mathbf{E}}_z^T)$. Specifically, since both vectors contain the same elements there exists a unique permutation matrix \mathbf{P} such that $\text{vec}(\tilde{\mathbf{E}}_z^T) = \mathbf{P}\text{vec}(\mathbf{E}_z)$ with $\mathbf{P}^{-1} = \mathbf{P}^T$. Using this result in Eq. (12), we have $\text{vec}(\mathbf{A}_z^T) = \tilde{\mathbf{G}}\tilde{\mathbf{X}}\text{vec}(\tilde{\mathbf{E}}_z^T)$, where

$$\tilde{\mathbf{G}} = \mathbf{P}\mathbf{G}\mathbf{P}^T = \begin{pmatrix} \tilde{\mathbf{G}}^{(1,1)} & \tilde{\mathbf{G}}^{(1,2)} & \dots & \tilde{\mathbf{G}}^{(1,M)} \\ \tilde{\mathbf{G}}^{(2,1)} & \tilde{\mathbf{G}}^{(2,2)} & \dots & \tilde{\mathbf{G}}^{(2,M)} \\ \vdots & \vdots & \dots & \vdots \\ \tilde{\mathbf{G}}^{(M,1)} & \tilde{\mathbf{G}}^{(M,2)} & \dots & \tilde{\mathbf{G}}^{(M,M)} \end{pmatrix} \quad (15)$$

and the elements of the block matrices are given by

$$\tilde{\mathbf{G}}^{(m,i)} \Big|_{nj} = (k_b \delta x)(k_b \delta y) G^w(\mathbf{x}_{mn} - \mathbf{x}_{ij}). \quad (16)$$

Furthermore, the diagonal contrast matrix $\tilde{\mathbf{X}}$ is given by $\tilde{\mathbf{X}} = \mathbf{P}\mathbf{X}\mathbf{P}^T = \text{diag}(\chi_{1,:}) \oplus \text{diag}(\chi_{2,:}) \oplus \dots \oplus \text{diag}(\chi_{M,:})$.

By pre- and postmultiplying matrix \mathbf{G} by the permutation matrix \mathbf{P} and its transpose, respectively, the permutation matrix turns the N -by- N block matrix \mathbf{G} with blocks of order M into an M -by- M block matrix $\tilde{\mathbf{G}}$ with blocks of order N . In other words, the roles of the indices (n, j) and (m, i) are interchanged, that is, in matrix $\tilde{\mathbf{G}}$ the indices (m, i) label the blocks and the elements within a block are labeled by the (n, j) indices, while for matrix \mathbf{G} it is exactly the other way around.

The preconditioner for two-dimensional problems is now introduced as $\mathbf{M} = \mathbf{I} - \mathbf{C}$, where \mathbf{C} has the same block partitioning as matrix $\mathbf{G}\mathbf{X}$, that is, matrix \mathbf{C} is of the form

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}^{(1,1)} & \mathbf{C}^{(1,2)} & \dots & \mathbf{C}^{(1,N)} \\ \mathbf{C}^{(2,1)} & \mathbf{C}^{(2,2)} & \dots & \mathbf{C}^{(2,N)} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{C}^{(N,1)} & \mathbf{C}^{(N,2)} & \dots & \mathbf{C}^{(N,N)} \end{pmatrix} \quad (17)$$

and each block matrix $\mathbf{C}^{(n,j)}$ is taken to be a circulant matrix of order M . The blocks are determined as in the one-dimensional case, that is, block matrix $\mathbf{C}^{(n,j)}$ is uniquely determined by

$$\mathbf{C}^{(n,j)} = \underset{\mathbf{Z} \text{ circulant}}{\text{argmin}} \|\mathbf{G}^{(n,j)} \text{diag}(\chi_{:,j}) - \mathbf{Z}\|_{\text{F}}. \quad (18)$$

The elements of each block can be computed explicitly using Eq. (4).

Now that we have obtained matrix \mathbf{C} , we still need an efficient way of computing the action of

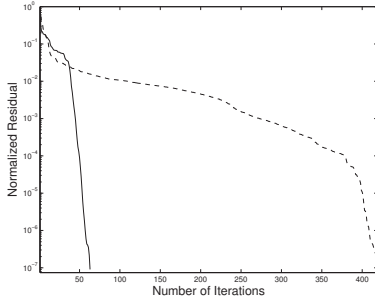


Figure 1: Convergence history of preconditioned (solid line) and unpreconditioned (dashed line) GMRES for a 1.6λ square block with a contrast $\chi = 32$.

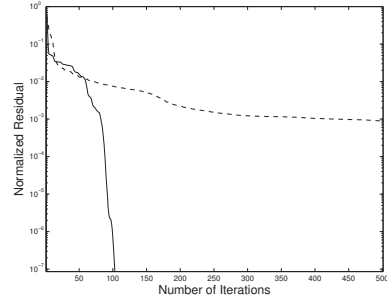


Figure 2: Convergence history of preconditioned (solid line) and unpreconditioned (dashed line) GMRES for a 1.6λ square block with a contrast $\chi = 64$.

\mathbf{M}^{-1} on a vector. First, observe that all circulant matrices $\mathbf{C}^{(n,j)}$ are diagonalized by the same DFT matrix \mathbf{F} , that is, we have $\mathbf{C}^{(n,j)} = \mathbf{F}^H \mathbf{D}^{(n,j)} \mathbf{F}$ for $n, j = 1, 2, \dots, N$. This allows us to write

$$\mathbf{C} = [\mathbf{P}(\mathbf{I}_N \otimes \mathbf{F})]^H \begin{pmatrix} \tilde{\mathbf{D}}^{(1,1)} & \tilde{\mathbf{D}}^{(1,2)} & \dots & \tilde{\mathbf{D}}^{(1,M)} \\ \tilde{\mathbf{D}}^{(2,1)} & \tilde{\mathbf{D}}^{(2,2)} & \dots & \tilde{\mathbf{D}}^{(2,M)} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{D}}^{(M,1)} & \tilde{\mathbf{D}}^{(M,2)} & \dots & \tilde{\mathbf{D}}^{(M,M)} \end{pmatrix} [\mathbf{P}(\mathbf{I}_N \otimes \mathbf{F})],$$

where $\tilde{\mathbf{D}}^{(m,i)}|_{n,j} = \mathbf{D}^{(n,j)}|_{m,i}$. Since all $\mathbf{D}^{(n,j)}$ are diagonal, we have $\tilde{\mathbf{D}}^{(m,i)} = \mathbf{0}$ if $m \neq i$ and the diagonal blocks are given by

$$\tilde{\mathbf{D}}^{(m,m)}|_{n,j} = \mathbf{d} \left(\mathbf{F} \mathbf{G}^{(n,j)} \text{diag}(\chi_{:,j}) \mathbf{F}^H \right) |_{m,m}.$$

Notice that $\tilde{\mathbf{D}}^{(m,m)}$ is a Toeplitz matrix if the contrast is constant along the j th column ($\chi_{:,j} = \chi$). If the contrast $\chi_{:,j}$ is not constant then we take the arithmetic average along the j th column (cf. Eq. (4) for $i = 0$).

With the above results, we now have for the inverse of our preconditioner

$$\mathbf{M}^{-1} = [\mathbf{P}(\mathbf{I}_N \otimes \mathbf{F})]^H \begin{pmatrix} \mathbf{M}_{1,1}^{-1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{2,2}^{-1} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{M}_{M,M}^{-1} \end{pmatrix} [\mathbf{P}(\mathbf{I}_N \otimes \mathbf{F})].$$

with $\mathbf{M}_{m,m} = \mathbf{I} - \tilde{\mathbf{D}}^{(m,m)}$ for $m = 1, 2, \dots, M$. We observe that to compute the action of \mathbf{M}^{-1} on a vector, we need to solve M one-dimensional Toeplitz systems (apart from the Fourier transform and permutation operation, of course). These systems can again be solved iteratively using the op-

timal circulant preconditioning techniques of Section 2.

4 NUMERICAL EXAMPLES

To illustrate the performance of the preconditioner, we consider a homogeneous square block embedded in a homogeneous background medium. The side length of the block is 1.6λ , where λ is the wavelength in the background medium. Figure 1 shows the convergence history of GMRES in case the block has a contrast $\chi = 32$. For this example, preconditioned GMRES requires 63 iterations to reach the specified tolerance of 10^{-7} , while unpreconditioned GMRES requires 419 iterations. On an Intel Core i7 Macbook Air, it took preconditioned GMRES 0.8 minutes to reach the desired tolerance level, while unpreconditioned GMRES required 3.7 minutes. Finally, in Fig. 2 we show the convergence history of preconditioned and unpreconditioned GMRES for a contrast $\chi = 64$. Preconditioned GMRES reaches the tolerance level within 103 iterations (computation time: 1.6 minutes), while unpreconditioned GMRES only converges to a level of about $8.9 \cdot 10^{-4}$ within the maximum number of 500 iterations.

References

- [1] R. H. F. Chan and X. Q. Jin, “An introduction to iterative Toeplitz solvers,” SIAM, Philadelphia, 2007.
- [2] T. Chan, “An Optimal Circulant Preconditioner for Toeplitz Systems,” *SIAM J. Sci. Stat. Comput.*, Vol. 9, 1998, pp. 766 – 771.
- [3] R. F. Remis, “Circulant preconditioners for domain integral equations in electromagnetics,” *Conf. Proc. ICEAA12*, pp. 337 – 340, 2012.