EFFECT OF CHOICE OF BOUNDARY CONDITION ON THE NUMERICAL EFFICIENCY OF DIRECT SOLUTIONS OF FINITE DIFFERENCE FREQUENCY DOMAIN SYSTEMS WITH PERFECTLY MATCHED LAYERS

Nathan Zhao  
Department of Applied Physics  
Stanford University  
Stanford, CA 94305  
nzz2102@stanford.edu

Shanhui Fan  
Department of Electrical Engineering  
Stanford University  
Stanford, CA 94305  
shanhui@stanford.edu

March 1, 2022

ABSTRACT

Direct solvers are a common method for solving finite difference frequency domain (FDFD) systems that arise in numerical solutions of Maxwell’s equations. In a direct solver, one factorizes the system matrix. Since the system matrix is typically very sparse, the fill-in of these factors is the single most important computational consideration in terms of time complexity and memory considerations. As a result, it is of great interest to determine ways in which this fill-in can be systematically reduced. In this paper, we show that in the context of commonly used perfectly matched boundary layer methods, the choice of boundary condition behind the perfectly matched boundary layer can be exploited to reduce fill-in incurred during the factorization, leading to significant gains of up to 40% in the efficiency of the factorization procedure. We illustrate our findings by solving linear systems and eigenvalue problems associated with the FDFD method.

1 Introduction

In any finite difference simulation, the domain must be truncated and a boundary condition is imposed. Common boundaries include the periodic boundary condition or the Dirichlet boundary condition. However, in many applications, one wants to simulate an "infinite" domain in the sense that any wave incident on the boundary should have no reflection. In one dimension, it is possible to eliminate all reflections via the Mur absorbing boundary condition [1]. However in $n > 2$ dimensions, there is no absorbing boundary condition that can eliminate reflections for waves incident from all angles. Instead, it is common to surround the computational domain by perfectly matched layers (PML), a specially designed anisotropic absorber [1,2]. The PML region is typically truncated using a boundary condition such as the periodic or the Dirichlet boundary condition.

The PML has been extensively used in computational electromagnetics, for both frequency and time-domain solutions of the Maxwell’s equations [1,5]. In the time domain, significant efforts have been invested in reducing the reflection for a given number of layers in the PML region [6,8]. Various formalisms of the PML have also been developed to ensure strong suppression of reflection over broad bandwidth [2,9]. In the frequency domain, there are additional considerations in the applications of PML. For example, Ref. [10] has shown that the condition number of the system matrix, and hence the performance of iterative solutions of Maxwell’s equations in the frequency domain, can vary drastically for different formalisms of PML, even though these formalisms offer similar wall-clock performance in time-domain algorithms.

In this paper we provide a discussion of PML in the context of direct solver for Maxwell’s equations in the frequency domain. In such direct solvers, one typically performs an LU decomposition of the system matrix. We show that the fill-in in the LU decomposition strongly depends on the boundary condition that is used to truncate the PML. In
particular, the use of the Dirichlet boundary condition leads to far less fill-in and hence significant efficiency gain, as compared with the use of the periodic boundary condition. This result is in contrast with both time-domain algorithms, as well as with iterative algorithms in the frequency domain for the solutions of Maxwell’s equations, where the use of these two boundary conditions do not significantly affect the performance of the algorithms. Our result is important for the development of direct solvers for Maxwell’s equations in the frequency domain.

2 A Brief Review of the Finite Difference Frequency Domain Method

2.1 Linear system formulation

The Maxwell’s equation written for the $\mathbf{H}$-field is:

$$\nabla \times \frac{1}{\epsilon_r(r)} \nabla \times \mathbf{H}(r) - \omega^2 \mu_0 \mathbf{H}(r) = -i\omega \mathbf{M}(r) \quad (1)$$

where $\epsilon_r$ denotes a three-dimensional dielectric tensor, $\omega$ is the frequency, $\mathbf{H}(r)$ denotes the vector magnetic field, $\mathbf{M}(r)$ is the vector magnetic source, and the magnetic permeability is assumed to be $\mu_0$ in the entire computational domain. In the finite-difference frequency-domain (FDFD) method, one discretizes Eq. (1) via finite differences on a Yee grid [2][11][13] to obtain a system of linear equations:

$$\hat{A} \mathbf{x} = \mathbf{b} \quad (2)$$

The matrix $\hat{A}$ corresponds to the left hand side of Eq. (1) and $\mathbf{x}$ is a vector corresponding to the discretized version of $\mathbf{H}$. $\mathbf{b}$ corresponds to the source term on the right hand side.

In order to obtain a solution to Eq. (2), a number of algorithms exist, such as Gaussian elimination [14]. However, a common method is to factorize the matrix $\hat{A}$ into $\hat{A} = \hat{L}\hat{U}$, where $\hat{L}$ is a lower triangular matrix and $\hat{U}$ is an upper triangular matrix. The solution can then be obtained:

$$\mathbf{x} = \hat{U}^{-1}(\hat{L}^{-1}\mathbf{b}) \quad (3)$$

Eq. (3) is typically called back-substitution. The computation of $L^{-1}b$ and $U^{-1}(L^{-1}b)$ is cheaper as compared with the full Gaussian elimination of the system matrix $\hat{A}$ [15][16]. However, the process of determining the factorization requires a modified version of Gaussian elimination, so the overall time complexity of using the LU method is the same as the Gaussian elimination [14]. On the other hand, if one has multiple right hand sides to solve, then the LU factorization becomes efficient as the $L$ and $U$ factors can be re-used.

2.2 Linear Eigenvalue System Formulation

The LU decomposition is also useful for solving eigenvalue problems in electromagnetics, an example of which is:

$$\nabla \times \frac{1}{\epsilon_r(r)} \nabla \times \mathbf{H}(r) = \omega^2 \mu_0 \mathbf{H}(r) \quad (4)$$

Following the same discretization procedure as that leading to Eq. (7), we obtain the eigenvalue problem:

$$\hat{A}\mathbf{x} = \lambda \mathbf{x} \quad (5)$$

where $\hat{A}$ now corresponds to the left-hand side of Eq. (5) and $\lambda = \mu_0\omega^2$. Eigenvalues are typically determined iteratively by using the Arnoldi iteration and its variants where one repeatedly applies the matrix $\hat{A}$ to a test vector [17][18]. However, direct application of these methods can only produce the largest $k$ eigenvalues of a system. In computational electromagnetics, one typically is interested instead in eigenvalues that are near a target value $\sigma$. For such a problem, a reformulation of the eigenvalue problem is required [17][19]. Instead of solving Eq. (5) we solve:

$$(\hat{A} - \sigma I)^{-1}\mathbf{x} = \beta \mathbf{x} \quad (6)$$

where $I$ represents the identity matrix. The eigenvalues $\beta$ can be related to eigenvalues of $A$ via: $\beta = 1/(\lambda - \sigma)$. Thus the eigenvalues of $\hat{A}$ that are closest to $\sigma$ correspond to the largest eigenvalues of $(\hat{A} - \sigma I)^{-1}$. In applying iterative methods to obtain the largest eigenvalues of $(\hat{A} - \sigma I)^{-1}$, we need to repeatedly apply $(\hat{A} - \sigma I)^{-1}$. In the FDFD method, $\hat{A}$ is sparse, but $(\hat{A} - \sigma I)^{-1}$ is fully dense. For large problems it is too costly memory-wise to form $(\hat{A} - \sigma I)^{-1}$ explicitly. The efficient way to do this is to compute an LU factorization of $\hat{A} - \sigma I$ [20]. Thus, the LU factorization is important for the eigenvalue problems that arise in the FDFD method as well.
2.3 Brief Review of Boundary Conditions

In the FDFD method the computational domain is finite and a boundary condition needs to be applied on the surface of the computational domain. This is illustrated in Fig. 1 where we show an undirected graph representation of a simple square grid consisting of \( N = N_x \times N_y \) nodes with \( N_x = N_y = 10 \). The points at which the fields are evaluated are termed nodes. We denote the fields evaluated at the nodes as \( x(i,j) \), where \( i \) is the row index and \( j \) is the column index. For a 2D problem, the Maxwell’s equations can be written in terms of only \( E_z \) or \( H_z \) fields, where \( z \) is the out-of-plane direction, hence \( x \) can correspond either to the \( E_z \) or the \( H_z \) fields. The outermost nodes are the boundary nodes where the boundary condition is enforced. The edges represent the node-to-node coupling as derived from finite-difference approximation of Eq. (1). In Fig. 1a, we show the periodic boundary, which can be described as: \( x(N_x,j) = x(0,j) \) and \( x(i,N_y) = x(i,0) \). In Fig. 1b, we show a Dirichlet boundary condition: \( x(N_x,j) = x(0,j) = x(i,N_y) = x(i,0) = 0 \). When \( x \) corresponds to the \( H(E) \) fields, the Dirichlet boundary condition corresponds to the perfectly magnetic (electric) conductor boundary condition.

![Figure 1: Two dimensional finite difference grid with a five point stencil. a) A grid truncated with the periodic boundary condition. The boundary connections which connect left boundary to right and top to bottom are shown as trailing black edges. b) Same grid with the Dirichlet boundary condition.](image)

In typical FDFD simulations one often surrounds a computation domain of interest with perfectly match layers (PML) [1, 10]. These layers are used to ensure that any wave exiting the computational domain is nearly perfectly absorbed, which is important for simulating an open electromagnetic system. When the PML are used, the entire computational domain consists of the domain of interest and the PML regions. Outside the PML regions, the computational domain still needs to be truncated. Either the periodic or the Dirichlet boundary conditions can be used for this purpose.

The choice of either the periodic or the Dirichlet boundary conditions outside the PML regions does not affect their absorption properties. Moreover, in the finite difference time domain method, or when iterative solvers are used in the finite difference frequency domain method, these two boundary conditions result in very similar computational cost. Therefore, there has been very little attention paid to the choice between these two boundary conditions. As we will show in this paper, however, this choice in fact makes a significant difference when direct solvers are used in the finite difference frequency domain method, since the choice significantly influences the fill-in behaviors in the LU decomposition of the system matrix.

3 A Brief Review of the Fill-in of Sparse Linear Systems

In general, \( \hat{A} \) for Eq. (2) or Eq. (5) is sparse. A key consideration in the LU factorization is the issue of fill-in, which is the accumulation of additional non-zero elements in \( \hat{L} \) and \( \hat{U} \) outside those of \( \hat{A} \). In general, fill-in during the factorization process for a sparse linear system depends sensitively on the detailed sparsity pattern of the system matrix.
Consider, for example, the following factorization of a matrix:

\[
\begin{bmatrix}
p & p & p & p \\
p & p & p & p \\
p & p & p & p \\
p & p & p & p \\
p & p & p & p \\
\end{bmatrix} =
\begin{bmatrix}
p \\
p & p & p \\
p & p & p & p \\
p & p & p & p & p \\
p & p & p & p & p \\
\end{bmatrix} \begin{bmatrix}
1 & p & p & p & p \\
1 & p & p \\
1 & p \\
1 \\
1 \\
\end{bmatrix}
\] (7)

where \( p \) denotes a matrix element which is nonzero (element values may be different). Because the first row has no zeros, the factorization process results in \( \hat{L} \) and \( \hat{U} \) factors that are dense [14]. In contrast, consider the factorization of an equivalent system as obtained from the left hand size of Eq. (7) by permuting the row index:

\[
\begin{bmatrix}
p & p & p & p \\
p & p & p & p \\
p & p & p & p \\
p & p & p & p & p \\
p & p & p & p & p \\
\end{bmatrix} =
\begin{bmatrix}
p & p & p \\
p & p & p & p \\
p & p & p & p & p \\
p & p & p & p & p \\
p & p & p & p & p \\
\end{bmatrix} \begin{bmatrix}
1 & p & p & p & p \\
1 & p \\
1 & p \\
1 & p \\
1 \\
\end{bmatrix}
\] (8)

The \( \hat{L} \) and \( \hat{U} \) factors resulting from Gaussian elimination do not have any fill-in. In general, the order of the equations as well as the order of the variables can drastically change the fill-in. The operations which generate these reorderings can be expressed as preconditioners on \( A \) via a set of permutation matrices \( \hat{P} \) and \( \hat{Q} \): \( A \rightarrow \hat{P} \hat{A} \hat{Q} \), where \( \hat{P} \) performs row-wise and \( \hat{Q} \) performs column-wise permutations. Unfortunately, the problem of finding the optimal permutation to produce the least amount of fill-in is NP-complete [21], which means there are no known, at-worst polynomial time algorithms to determine the optimal permutation. Instead, there is a wide array of different heuristic or approximate methods to reduce fill-in.

The state-of-the-art re-orderings are based on the exact minimum degree algorithm and their variants, such as the methods of approximate minimum degree (AMD) and nested dissection [22–26]. All of these methods effectively rely on the minimum degree heuristic, which is the observation that first eliminating variables in the factorization process with fewer couplings (or nonzeros in the row/column) tends to yield less fill-in overall.

4 Reducing Fill-in by Reducing Connectivity within the PML

As discussed in Section 2.3, the choice of boundary condition backing a PML can vary. As a result, we have the freedom to choose the boundary condition behind the PML. In this section, we now show how the choice of boundary condition behind the PML can affect the efficiency of sparse direct solvers, specifically impacting the fill-in during factorization.

From Figure 1, we can apply the minimum degree heuristic by counting the neighbors of the individual nodes on the grid. The number of neighbors corresponds to the number of nonzero elements in each row of \( A \). For the periodic boundary case shown in Fig. [1], all nodes including the boundary have 4 neighbors. By the minimum degree heuristic, there is no preference in picking which rows to eliminate first. Meanwhile, for the Dirichlet boundary in Fig. [1b], the four corner nodes have 2 neighbors, the edge nodes have 3 neighbors and then all other nodes have 4 neighbors. By the minimum degree heuristic in this case, we can see that the rows corresponding to the boundary nodes should be eliminated first.

Finally, we can take advantage of the fact that the nodes adjacent to the Dirichlet boundary have fields that are generally close to zero, since the fields have been strongly by the PML. As a result, we can also further modify the connectivity to minimize the number of neighbors per node, as we show in Fig. [2].
Figure 2: a) Optimized grid with reduced connectivity of the boundary. b) Same as a) with the nodes color coded by their degree (deg) in the graph representation.

### Table 1: Changes in the number of couplings in a grid of $N$ nodes for different boundary conditions. The 2nd column shows the analytic expression for the number of total couplings in the grid. The third column shows the values of these expressions for the specific case of $N = 10000$.

| Boundary Condition of Grid | Total Grid Couplings for $(N_x = N_y)$ | $N_x = 100$ |
|----------------------------|----------------------------------------|-------------|
| Periodic                  | $4N_x^2$                               | 40000       |
| Dirichlet                 | $4(N_x - 2)^2 + 4(N_x - 2) + 8$        | 38816       |
| Modified Dirichlet        | $4(N_x - 2)^2 + N_x - 2$               | 38514       |

In Fig. 3, we show the influence of the boundary conditions on the fill-in behavior for the small $10 \times 10$ grids shown in Figs. 1 and 2. The use of such a small grid allows easier visualization of various quantities involved. We solve a 2D problem corresponding to setting all terms involving $\partial_z$ to zero in Eq. (1). We simulate a magnetic source in vacuum for the magnetic field perpendicular $H_z$ to the 2D plane:

$$
\left( \frac{\partial}{\partial x} \epsilon_r(x, y) \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \epsilon_r(x, y) \frac{\partial}{\partial y} + \mu_0 \omega^2 \right) H_z = i\omega M_z
$$

(9)

where we assume $\epsilon_r$ is isotropic, $M_z$ is the $z$-component of the magnetic source current. We consider the periodic boundary conditions, the Dirichlet boundary conditions, and the modified Dirichlet boundary conditions as discussed above where the connectivity at the boundary is reduced. The first column of Figure 3 shows the sparsity pattern of the system matrices for the three boundary conditions. The sparsity patterns of the three matrices are similar since they differ only on the boundary. Compared with the system matrix from the Dirichlet boundary condition (Fig. 3d), the matrix from the periodic boundary condition (Fig. 3a) has extra non-zero elements at the lower-left and upper-half corners in the figure due to the long-range coupling induced by the periodic boundary condition. The matrix from the modified Dirichlet (Fig. 3g) boundary condition has less non-zero elements near the diagonal as results from the modifications.

The second column in Figure 3 shows the $\hat{L}$ factor as obtained by directly factoring the corresponding matrices in the first column as is without any reordering. For the case with periodic boundary condition (Fig. 3b), the $\hat{L}$ factor fills in near the diagonal and the lower edge in the figure. The fill-in at the lower edge arises from the long-range coupling as induced by the periodic boundary condition. By switching to the Dirichlet boundary condition (Fig. 3e), the fill-in near the lower edge disappears and the number of non-zero elements in the $\hat{L}$ factor significantly decreases. Modifying the Dirichlet boundary condition further reduces the fill-in and the number of non-zero elements (Fig. 3h).

The third column in Figure 3 shows the $\hat{L}$ factor as obtained by factoring the corresponding matrices in the first column using a minimum degree heuristics to reorder the matrix [22]. Compared with the second column, the use of the minimum degree heuristics significantly reduces the number of fill-in for all three cases. Also, the number of non-zero elements reduces as we go from the case with the periodic boundary condition (Fig. 3e), to the case with the Dirichlet
boundary condition (Fig. 3f), to the case with the modified Dirichlet boundary condition (Fig. 3i). The $\hat{L}$ factor shown in Fig. 3i, which was obtained for the modified Dirichlet boundary condition using the minimum degree heuristics, has 456 non-zero elements, as compared with the number of non-zero elements of 500 in the system matrix with the periodic boundary condition. The result in Figure 3 provides a direct demonstration that the use of the modified Dirichlet boundary condition, in combination with the minimum degree heuristics, can significantly reduce the fill-ins.

Figure 3: In the first column of Fig. 3, we show the sparsity pattern of $\hat{A}$ corresponding to the finite difference discretization of the $10 \times 10$ grid with the a) periodic, d) Dirichlet boundary, and g) modified Dirichlet boundary with reduced connectivity at the boundary respectively. In the second column, we show the corresponding $\hat{L}$ factors for factorizing these three matrices as is. In the third column, we show corresponding $\hat{L}$ factors for factorizing these three matrices using a minimum degree heuristic to reorder the matrices. The number of nonzero elements (nnz) for each matrix is also provided in each figure.

5 Numerical Experiments

In this section, we show that the concept as described above can be applied to typical simulations of electromagnetic waves. We consider both the driven case with a source, and eigenmode calculations without a source. For both cases, we show that choosing either the periodic or the Dirichlet boundary conditions behind the PML does not affect the accuracy of the results, but the choice of the Dirichlet boundary condition leads to better numerical performance.

5.1 Driven Problem

We solve Eq. (9) as in the previous section but now the interior domain consists of $N_x \times N_y$ with $N_x = N_y = 301$ grid cells surrounded with a PML with a thickness of 30 cells to ensure minimal reflection [27]. The domain has a length $a = 4\lambda$ where $\lambda$ is the wavelength of the source.

First, we confirm that the use of either the periodic boundary or the modified Dirichlet boundary in conjunction with a PML indeed has no effects on the solution within the interior domain. In Fig. 4 we show the result where the
interior domain consists of vacuum and we put a point source at the center of the interior domain. For the two boundary conditions, the field distributions inside the computational domain are almost identical, as can be seen by comparing visually the field distributions shown in Fig. 4a and b, or by examining the field distributions along a line in the computational domain as shown in Figure 4c.

For the matrix \( \hat{A} \) corresponding to the driven problem, we show in Fig. 5 the difference in factorizing the matrix with the periodic boundary vs the Dirichlet boundary. In Fig. 5a and b, we show the fill-in of the \( \hat{L} \) factor with the modified Dirichlet and periodic boundary respectively for the system shown in Figure 4. The \( \hat{L} \) factor from the modified Dirichlet boundary is noticeably sparser than the \( \hat{L} \) corresponding to the periodic boundary.

In Fig. 5, we vary the size the domain by keeping the domain to be a square shape and varying \( N_x = N_y \). We plot the percentage reduction in the number of non-zero elements in the \( \hat{L} \) factor as we change from the periodic to the modified Dirichlet boundary condition. The percentage reduction fluctuates significantly as \( N_x \) varies, since the reordering strategies are heuristic. But in general, we see significant reduction for all \( N_x \)’s. On average the percentage reduction decreases as \( N_x \) increases, which would lead to a smaller effect on the relative fill-in. What is remarkable is that the boundary couplings clearly have a disproportionately large effect on the fill-in of \( \hat{L} \) factor, even though the boundary nodes are often just 1% or less of the total nodes in the entire grid. Even with \( N_x \approx 1000 \), there are cases where the use of Dirichlet boundary condition results in 33% fewer nonzeros in the \( \hat{L} \) as compared with the use of the periodic boundary condition.
5.2 Eigenvalue Problem

As a second example, we consider an eigenvalue problem that commonly arises in the study of nanophotonic waveguides. In this problem, we consider a structure that is periodic along the $x$-direction. For a given operating frequency $\omega$, we solve for the propagating wavevector $k_x$. To develop the formalism, we substitute via the Bloch theorem $H_z(x, y) \rightarrow H_z(x, y)e^{ik_xx}$ in Eq. (9) to obtain:

$$\frac{1}{\mu_0} \left( \frac{1}{\epsilon_r} \partial_x + \partial_y \left( \frac{1}{\epsilon_r} \right) \partial_y + \omega^2 \mu_0 + i k_x \left( \partial_x \epsilon_r^{-1} + \epsilon_r^{-1} \partial_x \right) \right) H_z e^{ik_xx} = \frac{1}{\mu_0 \epsilon_r} k_x^2 H_z e^{ik_xx} \tag{10}$$

which is a quadratic eigenvalue problem of the form:

$$\hat{M} \mathbf{v} + \lambda^2 \hat{K} \mathbf{v} + \lambda \hat{C} \mathbf{v} = 0 \tag{11}$$

where $\lambda = k_x$ and:

$$\hat{K} = \frac{1}{\mu_0} \epsilon_r \tag{12}$$

$$\hat{M} = \frac{1}{\mu_0} \left( \frac{1}{\epsilon_r} \partial_x + \partial_y \left( \frac{1}{\epsilon_r} \right) \partial_y + \omega^2 \right) \tag{13}$$

$$\hat{C} = \frac{i}{\mu_0} \left( \partial_x \epsilon_r^{-1} + \epsilon_r^{-1} \partial_x \right) \tag{14}$$

Eq. (11) can be reformulated into an equivalent linear, generalized eigenvalue problem of the form:

$$\hat{D} \lambda \mathbf{w} = \hat{B} \mathbf{w} \tag{15}$$

where $\mathbf{w} = [x, \lambda x]$ with $x$ denoting the vector formed by $H_z(x, y)$ and where:

$$\hat{D} = \begin{bmatrix} \hat{K} & 0 \\ 0 & \hat{I} \end{bmatrix} \tag{16}$$

and

$$\hat{B} = \begin{bmatrix} \hat{C} & \hat{M} \\ -\hat{I} & 0 \end{bmatrix} \tag{17}$$

After discretization using a computational domain consisting of a single unit cell of the waveguide structure, $\hat{0}$ denotes a matrix of zeros of dimension $N_x N_y$. In the waveguide eigenvalue problem, similar to the discussion in Section (2.2), we are typically interested in modes with wavevector near a particular value $\sigma$. Therefore, following the same derivation as in Section (2.2), the matrix that needs to factorized is $\hat{B} - \sigma \hat{D}$.

As an illustration we consider the waveguide structure schematically shown in Figure 6a. The permittivity of the materials used is provided in the caption of Figure 6a. The waveguide consists of two walls. Each wall consists of a mixture of dielectric and a Drude metal. The thickness of each wall is 0.3 $\mu$m. The metal in the wall has a width along the $x$-direction of 0.04 $\mu$m. Two unit cells are shown in Figure 6a. By operating this system at $\lambda = 2 \mu$m, the walls become near-completely reflective, creating a hollow-core waveguide with light guided in the vacuum region between the walls.

To simulate this structure we use a computational domain consisting of a single unit cell shown in Figure 6a. We discretize the computational domain with $N_x \times N_y$ grid cells where $N_x = 100$ and $N_y = 150$. Along the $x$-direction a periodic boundary condition is imposed at the edges of the computational domain. Along the $y$-direction we place PML regions at both edges. The PML regions consist of 10 grid cells. Outside the PML region we truncate the grid with either the modified Dirichlet or the periodic boundary conditions.

First, we validate that the use of different boundary conditions backing the PML does not affect the guided modes of the system. In Fig. 6b and c, we compare the real part of the field profiles of the structure for a sample mode with the same solved $k_x$ eigenvalue but with either the Dirichlet or periodic boundary conditions. We see an excellent agreement between the modal profiles. In Fig. 7a we compare the bandstructure of the waveguide with the periodic and Dirichlet boundary conditions backing the PML in the $y$-direction. To generate these bandstructures, we scan a discrete set of $\omega$ values, and look for the $k_x$ eigenvalues closest to 0. Additionally, due to spurious solutions that come from discretizing Eq. (10), a mode-filter must be implemented to produce the band-structure. Fig. 7a is the band structure with the Dirichlet boundary and Fig. 7b is the corresponding one with the periodic boundary. Again, we see that the band structures in both case are essentially identical, as expected.
A preprint - March 1, 2022

Figure 7: Band structures obtained with a) the modified Dirichlet boundary, b) with the periodic boundary condition. The blue and red dots correspond to the real and imaginary parts of the \( k_x \), respectively.

Figure 6: a) Schematic of the waveguide structure and the computational domain. The violet regions consist of a metal described by the Drude model with \( \epsilon(\omega) = 1 - \omega_p^2/(\omega_p^2 - i\gamma \omega) \), with \( \omega_p = 0.72\pi \times 10^{15} \) rad/s, \( \gamma = 5.5 \times 10^{12} \) s\(^{-1}\). The yellow regions consist of a dielectric material with a relative permittivity of 16. The blue region is vacuum. The gray regions are the PML regions. The structure is periodic along the \( x \)-direction. Plotted here are two unit cells of the structure. b) \( Re(H_z) \) modal profile of the structure with a Dirichlet boundary at the top and bottom edges of the computational domain at an operating frequency corresponding to a free space wavelength of 2 \( \mu \)m. b) The corresponding result with the periodic boundary at the top and bottom edges of the computational domain.

Now we compare the fill-in of factorizing \( \hat{B} - \sigma \hat{D} \) with the periodic versus the modified Dirichlet boundary conditions. In Fig. 8a-c we consider the case where \( \sigma = 0 \). In Fig. 8a, we show the sparsity pattern of the matrix \( \hat{B} \) in Eq. (17). In Fig. 8b and c, we show the fill-in of the \( \hat{L} \) factor of \( \hat{B} \) with the domain truncated by the periodic boundary and the modified Dirichlet boundary conditions, respectively. The fill-in patterns vary substantially between the two cases. Moreover, even though we have PML regions only along the \( y \) boundaries, we still achieve a 40\% reduction in the number of nonzeros. In the second row of Fig. 8, we show a case with \( \sigma \neq 0 \). Since \( \hat{B} \) and \( \hat{D} \) have different sparsity patterns, the sparsity pattern shown in Fig. 8d differ from that in Fig. 8a. In comparing Fig. 8e with Fig. 8f, we again see that the use of the modified Dirichlet boundary condition leads to less fill-in.

6 Conclusion

In summary, we consider direct solvers for the FDFD methods applied to Maxwell’s equations. We demonstrate that the choice of boundary condition used behind the PML can have significant effects on the numerical efficiency of the LU factorization of the system matrix. With the PML the physical properties of the solutions do not depend on the
choices of the boundary conditions. But we observe that the use of a modified Dirichlet boundary condition can lead to a reduction of fill-in by up to 40%, as compared to the commonly used periodic boundary conditions. Such a reduction translates to significant increase in computational speed and decrease in memory requirements for the direct solvers. While in our paper we focus on the PML, our results should be applicable for other numerical absorbers, such as the adiabatic absorbers [32]. Also, while we illustrate our observations with two-dimensional simulations, the same observations should be applicable for three-dimensional simulations as well. Our observation should also be relevant for finite-element frequency-domain solvers.

Our result is important for enhancing the performance of direct solvers in the FDFD methods. In addition, this result should have implications for several numerical tasks in computational electromagnetics that can benefit from efficient LU decomposition of the system matrix, such as adjoint variable method [12, 33], efficient line searches for gradient-based photonic optimization [34–36], and domain decomposition methods utilizing the Schur complement approach [37, 38].

References

[1] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185–200, oct 1994.
[2] Allen Taflove. *Computational Electrodynamics*. Artech House, Inc., USA, 3rd edition, 2005.
[3] J.-P. Berenger. Perfectly matched layer for the ffdtd solution of wave-structure interaction problems. *IEEE Transactions on Antennas and Propagation*, 44(1):110–117, 1996.
[4] Z.S. Sacks, D.M. Kingsland, R. Lee, and Jin-Fa Lee. A perfectly matched anisotropic absorber for use as an absorbing boundary condition. *IEEE Transactions on Antennas and Propagation*, 43(12):1460–1463, 1995.
[5] Steven G. Johnson. Notes on perfectly matched layers (pmls), 2021.
[6] F Collino and P B Monk. Optimizing the perfectly matched layer. *Computer Methods in Applied Mechanics and Engineering*, 164(1):157–171, 1998.
[7] Arti Agrawal and Anurag Sharma. Perfectly matched layer in numerical wave propagation: factors that affect its performance. *Appl. Opt.*, 43(21):4225–4231, Jul 2004.
[8] Ardavan Oskooi and Steven Johnson. Distinguishing correct from incorrect pml proposals and a corrected unsplit pml for anisotropic, dispersive media. *Journal of Computational Physics*, 230:2369–2377, 04 2011.

[9] J.-P. Berenger. Perfectly matched layer for the ftdtd solution of wave-structure interaction problems. *IEEE Transactions on Antennas and Propagation*, 44(1):110–117, 1996.

[10] Wonseok Shin and Shanhui Fan. Choice of the perfectly matched layer boundary condition for frequency-domain Maxwell’s equations solvers. *Journal of Computational Physics*, 231(8):3406–3431, apr 2012.

[11] Kane Yee. Numerical solution of initial boundary value problems involving maxwell’s equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, 1966.

[12] Georgios Veronis, Robert W. Dutton, and Shanhui Fan. Method for sensitivity analysis of photonic crystal devices. *Optics Letters*, 29(19):2288, oct 2004.

[13] Randall LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems (Classics in Applied Mathematics Classics in Applied Mathemat)*. Society for Industrial and Applied Mathematics, USA, 2007.

[14] G. Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, 2016.

[15] Xiaoye S. Li. An overview of SuperLU: Algorithms, implementation, and user interface, sep 2005.

[16] Timothy A. Davis. Algorithm 832: UMFPACK V4.3 - An unsymmetric-pattern multifrontal method. *ACM Transactions on Mathematical Software*, 30(2):196–199, jun 2004.

[17] R. B. Lehoucq, D. C. Sorensen, and C. Yang. Arpack users guide: Solution of large scale eigenvalue problems by implicitly restarted arnoldi methods., 1997.

[18] Gene H. Golub and Henk A. van der Vorst. Eigenvalue computation in the 20th century. *Journal of Computational and Applied Mathematics*, 123(1):35–65, 2000. Numerical Analysis 2000. Vol. III: Linear Algebra.

[19] Carmen Campos and Jose E Roman. Strategies for spectrum slicing based on restarted Lanczos methods. *Numerical Algorithms*, 60(2):279–295, 2012.

[20] Chang Peng and Daniel Boley. Methods for large sparse eigenvalue problems from waveguide analysis. In *Proceedings of the 1996 12th Annual Review of Progress in Applied Computational Electromagnetics*, pages 645–652, January 1996.

[21] Mihalis Yannakakis. Computing the minimum fill-in is np-complete. *SIAM Journal on Algebraic Discrete Methods*, 2(1):77–79, 1981.

[22] Patrick R. Amestoy, Timothy A. Davis, and Iain S. Duff. An approximate minimum degree ordering algorithm. *SIAM Journal on Matrix Analysis and Applications*, 17(4):886–905, 1996.

[23] Alan George. Nested dissection of a regular finite element mesh. *SIAM Journal on Numerical Analysis*, 10(2):345–363, 1973.

[24] John R Gilbert and Robert Endre Tarjan. The analysis of a nested dissection algorithm. *Numerische Mathematik*, 50(4):377–404, 1986.

[25] Pratik Agarwal and Edwin Olson. Variable reordering strategies for slam. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3844–3850, 2012.

[26] George Karypis and Vipin Kumar. MeTis: Unstructured Graph Partitioning and Sparse Matrix Ordering System, Version 4.0. [http://www.cs.umn.edu/~metis](http://www.cs.umn.edu/~metis), 2009.

[27] Allen Taflove, A. Oskooi, and S. Johnson. *Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology*. Artech House, 01 2013.

[28] Georgios Veronis and Shanhui Fan. Guided subwavelength plasmonic mode supported by a slot in a thin metal film. *Opt. Lett.*, 30(24):3359–3361, Dec 2005.

[29] Andre Nicolet and Christophe Geuzaine. Waveguide propagation modes and quadratic eigenvalue problems. In *6th International Conference on Computational Electromagnetics*, pages 1–3, 2006.

[30] Nathan Z. Zhao, Ian A. D. Williamson, Zhexin Zhao, Salim Boutami, and Shanhui Fan. Penetration depth reduction with plasmonic metalfilms. *ACS Photonics*, 6(8):2049–2055, 2019.

[31] Nathan Zhao, Ian A. D. Williamson, Zhexin Zhao, Salim Boutami, and Shanhui Fan. Penetration depth engineering in plasmonic metalfilms for enhanced reflection and confinement. In *Conference on Lasers and Electro-Optics*, page FF3E.7. Optica Publishing Group, 2020.

[32] Ardavan F. Oskooi, Lei Zhang, Yehuda Avniel, and Steven G. Johnson. The failure of perfectly matched layers, and towards their redemption by adiabatic absorbers. *Opt. Express*, 16(15):11376–11392, Jul 2008.
[33] Christopher M. Lalau-Keraly, Samarth Bhargava, Owen D. Miller, and Eli Yablonovitch. Adjoint shape optimization applied to electromagnetic design. *Optics Express*, 21(18):21693, sep 2013.

[34] Salim Boutami, Nathan Zhao, and Shanhui Fan. Large permittivity increments for efficient predictive photonic devices optimization. In Connie J. Chang-Hasnain, Andrei Faraon, and Weimin Zhou, editors, *High Contrast Metastructures IX*, volume 11290, pages 55–67. International Society for Optics and Photonics, SPIE, 2020.

[35] Salim Boutami, Nathan Zhao, and Shanhui Fan. Determining the optimal learning rate in gradient-based electromagnetic optimization using the shanks transformation in the lippmann–schwinger formalism. *Opt. Lett.*, 45(3):595–598, Feb 2020.

[36] Nathan Z. Zhao, Salim Boutami, and Shanhui Fan. Efficient method for accelerating line searches in adjoint optimization of photonic devices by combining schur complement domain decomposition and born series expansions. *Opt. Express*, 30(4):6413–6424, Feb 2022.

[37] Nathan Z. Zhao, Salim Boutami, and Shanhui Fan. Accelerating adjoint variable method based photonic optimization with Schur complement domain decomposition. *Optics Express*, 27(15):20711, jul 2019.

[38] Nathan Zhao, Sacha Verweij, Wonseok Shin, and Shanhui Fan. Accelerating convergence of an iterative solution of finite difference frequency domain problems via schur complement domain decomposition. *Opt. Express*, 26(13):16925–16939, Jun 2018.