
Eigendecompositions and fast eigensolvers for Maxwell equations

by So-Hsiang Chou^{*} and Wen-Wei Lin[†]

Maxwell Equations and Constitutive Relations

In this article we report some recent advances [7, 11] and their extensions in the computation of the band structures of photonic crystals, engineered periodic structures made of two or more materials. Our approach is to deduce explicit eigendecompositions for the discrete single and double curl operators associated with the eigenvalue problems resulting from the discretization of the Maxwell eigenvalue problem. These decompositions can be efficiently computed due to our specific and judicious design of the bases for the range and null spaces for which Fast Fourier Transform (FFT) can be used to accelerate the computation. Furthermore, as a result we can derive the null space free method to avoid the commonly annoying large null space issue when it comes to applying iterative eigenvalue solvers. The reduction of the problem size is often substantial. We will discuss a few signposts that made our approach possible and promising to tackle more complicated problems related to metamaterials.

To begin let us consider Maxwell equations in the most widely used form:

$$(1) \quad \begin{cases} \nabla \cdot D = \rho, \\ \nabla \cdot B = 0, \\ \nabla \times E = -\frac{\partial B}{\partial t}, \\ \nabla \times H = J + \frac{\partial D}{\partial t}, \end{cases}$$

^{*} Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403-0221, USA
E-mail: chou@bgsu.edu

[†] Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan
E-mail: wwlin@math.nctu.edu.tw

which mathematically summarizes the Gauss law, Gauss law for magnetism, Faraday's law of induction and Maxwell-Ampere's law. Here E denotes the electric field; H the magnetic field; D the electric displacement; B magnetic induction; ρ the electric charge density, and J the current density. The four basic equations must be supplied with constitutive relations relating the field variables in their working environment. For example, in vacuum state the constitutive relation is

$$(2) \quad \begin{aligned} D &= \epsilon_0 E, \\ B &= \mu_0 H, \end{aligned}$$

where ϵ_0 and μ_0 represent the vacuum permittivity and vacuum permeability, respectively. This classical form of Maxwell's equations has been studied and used extensively. However, more complicated constitutive relations are required in studying and developing modern materials. We need to consider generic relations that allow permittivity ϵ and permeability μ to be functions of frequency, electric and magnetic fields and space variables \mathbf{x} , i.e.,

$$(3) \quad \begin{aligned} D &= \epsilon(\omega, \mathbf{x}, E, H)E, \\ B &= \mu(\omega, \mathbf{x}, E, H)H. \end{aligned}$$

For example, constitutive relations in photonic crystals problems [21, 22] and in plasma phenomenon problems [13] fall into the above category. Furthermore, if the permeability and permittivity of medium are both negative, it was theoretically predicted in the sixties [20] that the phenomena of negative refraction and backward waves would occur. Nowadays [25] such phenomena can be observed in man-made bianisotropic complex materials. Their constitutive relations are characterized by the magnetoelectric as well

as ε (permittivity tensor) and μ (permeability tensor):

$$(4) \quad \begin{aligned} D &= \varepsilon E + \xi H, \\ B &= \mu H + \zeta E, \end{aligned}$$

where ξ and ζ are the coupling constants. Exploring Maxwell's equations under the above constitutive relations mathematically and numerically is one of the important endeavors in the research of metamaterials [16, 18, 19].

Continuous Maxwell Eigenvalue Problem

One of the major issues in studying the propagation of electromagnetic waves in a periodic structure is the determination of the band structure. In the context of photonic crystals, there may exist [23] a frequency range where no electromagnetic eigenmode exists. Such a range is termed a band gap, just as in solid state physics [12]. In such band gaps, the electromagnetic wave will be reflected, trapped and transported in the material. Mathematically, to find the gap we need to solve an eigenvalue problem associated with Maxwell equations subject to the associated constitutive relations. It is quite obvious that due to the complexity of the lattice structure of the material, the tensorial material parameters in the constitutive relation, and so on, numerical modeling turns out to be an indispensable and crucial tool in the investigation.

Before discussing the numerics, let us introduce some necessary assumptions or conditions to arrive at the eigenvalue problem for finding the band structure.

1. Time-Harmonic Assumption Generally, we assume that the source of the electromagnetic wave is harmonic in time:

$$E(\mathbf{x}, t) = e^{-i\omega t} E(\mathbf{x}) \quad \text{and} \quad H(\mathbf{x}, t) = e^{-i\omega t} H(\mathbf{x})$$

where ω represents the frequency and $\iota = \sqrt{-1}$. By this assumption, we can transform Maxwell equations from the time domain to the frequency domain. By convention, we denote the space field by the same symbol for the temporal-spatial field.

2. Bloch's Condition [3] If the electromagnetic wave travels in a periodic or a lattice structure with translation vectors $\mathbf{a}_l, l = 1, 2, 3$, then we will assume that the electric and magnetic fields $E(\mathbf{x}), H(\mathbf{x})$ satisfy the quasi-periodic condition

$$(5) \quad E(\mathbf{x} + \mathbf{a}_l) = e^{2\pi i \mathbf{k} \cdot \mathbf{a}_l} E(\mathbf{x}) \quad \text{and} \quad H(\mathbf{x} + \mathbf{a}_l) = e^{2\pi i \mathbf{k} \cdot \mathbf{a}_l} H(\mathbf{x}),$$

where the Bloch wave vector k lies in the first Brillouin zone. In the study of photonic crystals, three typical lattices in the cubic system are: the simple cubic (SC)

lattice, the body-centered cubic (BCC) lattice, and the face-centered cubic (FCC) lattice. They have different first Brillouin zones and the common angle between any two translation vectors of the lattice is 90° for SC, 120° for BCC and 60° for FCC. For simplicity of presentation we will only discuss the SC lattice case and provide references for other cases.

3. Boundary Conditions Sometimes, when an infinite medium is simulated in band structure calculations, perfect magnetic or perfect electric conductor boundary conditions (PMC or PEC) need to be imposed. Another commonly used boundary condition is the Perfect Matching Layer (PML) condition [2, 5, 9] when an infinite domain is reduced to a finite one.

Using the constitutive relations (4), time-harmonic assumption, and the source free condition $J = 0$ in (1), we can derive the following equation in matrix (operator) form

$$(6) \quad \begin{bmatrix} \nabla \times & 0 \\ 0 & \nabla \times \end{bmatrix} \begin{bmatrix} E \\ H \end{bmatrix} = i\omega \begin{bmatrix} \zeta & \mu \\ -\varepsilon & -\xi \end{bmatrix} \begin{bmatrix} E \\ H \end{bmatrix}.$$

Since ω is to be found as well, this equation represents a generalized eigenvalue problem (GEP) in operator form. This is the starting point for all the band structure problems discussed below.

Discrete Maxwell Eigenvalue Problem

Since it is hardly possible to solve the continuous GEP (6), we must develop a discrete model by proper discretization schemes that can preserve main properties such as the divergence conditions, i.e., Gauss law and Gauss law for magnetism in the continuous model. Below we discuss two main discretization techniques: Yee's staggered finite difference scheme and the finite element method using edge elements.

Eigendecomposition of Discrete Operators from Yee's Finite Difference Method

We use Yee's scheme [24] to discretize Eq. (6) on the cube of side length a , where a is the lattice constant. Let us find first the discrete versions of the single curl operators $\nabla \times E$ and $\nabla \times H$. Denoting $E = [E_1, E_2, E_3]^\top$, $H = [H_1, H_2, H_3]^\top$, and noting that

$$(7) \quad \begin{cases} \partial_y E_3 - \partial_z E_2 = (\nabla \times E)_1, \\ \partial_z E_1 - \partial_x E_3 = (\nabla \times E)_2, \\ \partial_x E_2 - \partial_y E_1 = (\nabla \times E)_3, \end{cases}$$

and

$$(8) \quad \begin{cases} \partial_y H_3 - \partial_z H_2 = (\nabla \times H)_1, \\ \partial_z H_1 - \partial_x H_3 = (\nabla \times H)_2, \\ \partial_x H_2 - \partial_y H_1 = (\nabla \times H)_3, \end{cases}$$

we discretize Eqs. (7) and (8) at the centers of cell faces and edges, respectively. More specifically, let δ_x , δ_y , and δ_z be the grid lengths along the x , y , and z directions respectively, and let n_1 , n_2 , and n_3 be the numbers of (computational) grid points in the x , y , and z directions, respectively, i.e., $n_1 = a/\delta_x$, $n_2 = a/\delta_y$, $n_3 = a/\delta_z$. The number of grid points is $n = n_1 n_2 n_3$.

The approximate function values due to finite differences are represented by the grid points indexed by i , j , and k and the ‘‘half grid points’’ indexed by $\hat{i} = i + \frac{1}{2}$, $\hat{j} = j + \frac{1}{2}$, and $\hat{k} = k + \frac{1}{2}$. An arbitrary point $(r\delta_x, s\delta_y, t\delta_z)$ in the computational domain is denoted by $\mathbf{x}(r, s, t)$, where $r, s, t \in \mathbb{R}$:

$$(9) \quad \mathbf{x}(r, s, t) = (r\delta_x, s\delta_y, t\delta_z).$$

For $i = 0, \dots, n_1 - 1$, $j = 0, \dots, n_2 - 1$, and $k = 0, \dots, n_3 - 1$, we use

- $E_1(\hat{i}, j, k)$, $E_2(i, \hat{j}, k)$, and $E_3(i, j, \hat{k})$ to denote approximate values of functions E_1 , E_2 , and E_3 in vector form, at the edge midpoints $\mathbf{x}(\hat{i}, j, k)$, $\mathbf{x}(i, \hat{j}, k)$, and $\mathbf{x}(i, j, \hat{k})$, respectively.
- $H_1(i, \hat{j}, \hat{k})$, $H_2(\hat{i}, j, \hat{k})$, and $H_3(\hat{i}, \hat{j}, k)$ denote approximate values of functions H_1 , H_2 , and H_3 in vector form at the central face points $\mathbf{x}(i, \hat{j}, \hat{k})$, $\mathbf{x}(\hat{i}, j, \hat{k})$, and $\mathbf{x}(\hat{i}, \hat{j}, k)$, respectively.

Finally, by using the vectorization function of a matrix $F \in \mathbb{C}^{m_1 \times m_2 \times m_3}$:

$$\text{vec}(F) = \begin{bmatrix} \text{vec}(F(1 : m_1, 1 : m_2, 1)) \\ \text{vec}(F(1 : m_1, 1 : m_2, 2)) \\ \vdots \\ \text{vec}(F(1 : m_1, 1 : m_2, m_3)) \end{bmatrix},$$

we define

$$(10) \quad \begin{aligned} \mathbf{e} &= [\mathbf{e}_1^\top \quad \mathbf{e}_2^\top \quad \mathbf{e}_3^\top]^\top \in \mathbb{C}^{3n}, \\ \mathbf{h} &= [\mathbf{h}_1^\top \quad \mathbf{h}_2^\top \quad \mathbf{h}_3^\top]^\top \in \mathbb{C}^{3n}, \end{aligned}$$

with $\mathbf{e}_\ell = \text{vec}(E_\ell)$ and $\mathbf{h}_\ell = \text{vec}(H_\ell)$, for $\ell = 1, 2, 3$, and the $3n$ -by- $3n$ diagonal matrix

$$B = \text{diag} \left[\text{vec}(B_1)^\top, \text{vec}(B_2)^\top, \text{vec}(B_3)^\top \right]^\top \in \mathbb{C}^{3n \times 3n}.$$

Let us just look at how to derive the matrix representation of the discretization for (7) by finite differences at the central face points $\mathbf{x}(i, \hat{j}, \hat{k})$, $\mathbf{x}(\hat{i}, j, \hat{k})$, and $\mathbf{x}(\hat{i}, \hat{j}, k)$, respectively. Since we are concentrating on SC lattices, the description will be easy. For other lattices the procedure is more involved but in any case, the key point is to explore the periodic properties associated with the lattice translation vectors.

Now the centered difference approximations of $\partial_x E_2$ at $\mathbf{x}(\hat{i}, \hat{j}, k)$ and $\partial_x E_3$ at $\mathbf{x}(\hat{i}, j, \hat{k})$ are

$$(11) \quad \frac{E_2(i+1, \hat{j}, k) - E_2(i, \hat{j}, k)}{\delta_x} \quad \text{and} \quad \frac{E_3(i+1, j, \hat{k}) - E_3(i, j, \hat{k})}{\delta_x},$$

respectively, for $i = 0, 1, \dots, n_1 - 1$, $j = 0, 1, \dots, n_2 - 1$, and $k = 0, 1, \dots, n_3 - 1$. Applying the periodic condition (5) along the lattice translation vector $\mathbf{a}_1 = a(1, 0, 0)^\top$, we have

$$\begin{aligned} E_2(n_1, \hat{j}, k) &= e^{i2\pi\mathbf{k}\cdot\mathbf{a}_1} E_2(0, \hat{j}, k) \quad \text{and} \\ E_3(n_1, j, \hat{k}) &= e^{i2\pi\mathbf{k}\cdot\mathbf{a}_1} E_3(0, j, \hat{k}), \end{aligned}$$

for $j = 0, 1, \dots, n_2 - 1$ and $k = 0, 1, \dots, n_3 - 1$. Thus, using the notation \mathbf{e}_ℓ defined in (10), the matrix representations of the discretizations in (11) are $C_1 \mathbf{e}_2$ and $C_1 \mathbf{e}_3$, respectively. Here,

$$C_1 = I_{n_2 \times n_3} \otimes K_1 \in \mathbb{C}^{n \times n}$$

and

$$K_1 = \frac{1}{\delta_x} \begin{bmatrix} -1 & 1 & & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ e^{i2\pi\mathbf{k}\cdot\mathbf{a}_1} & & & & -1 \end{bmatrix} \in \mathbb{C}^{n_1 \times n_1}.$$

Similar derivation can be done to the remaining two directions.

With the above motivation, we can now describe the expected form for the discrete GEP as:

$$(12) \quad \begin{bmatrix} C & 0 \\ 0 & C^* \end{bmatrix} \begin{bmatrix} \mathbf{e} \\ \mathbf{h} \end{bmatrix} = i\omega \begin{bmatrix} B_\zeta & D_\mu \\ -D_\varepsilon & -B_\xi \end{bmatrix} \begin{bmatrix} \mathbf{e} \\ \mathbf{h} \end{bmatrix},$$

where D_ε and D_μ are diagonal matrices, B_ζ and B_ξ are sparse matrices. (These matrices are the discretized versions of their continuous counterparts in (6).) In particular, C is the discrete matrix of curl operator, which can be defined by

$$(13) \quad C = \begin{bmatrix} 0 & -C_3 & C_2 \\ C_3 & 0 & -C_1 \\ -C_2 & C_1 & 0 \end{bmatrix} \in \mathbb{C}^{3n \times 3n},$$

with

$$(14a) \quad C_1 = \delta_x^{-1} (I_{n_3} \otimes I_{n_2} \otimes K_{\mathbf{a}_1, n_1}) \in \mathbb{C}^{n \times n},$$

$$(14b) \quad C_2 = \delta_y^{-1} (I_{n_3} \otimes K_{\mathbf{a}_2, n_2} \otimes I_{n_1}) \in \mathbb{C}^{n \times n},$$

$$(14c) \quad C_3 = \delta_z^{-1} (K_{\mathbf{a}_3, n_3} \otimes I_{n_2} \otimes I_{n_1}) \in \mathbb{C}^{n \times n},$$

in which

$$(15) \quad K_{\mathbf{a}, m} = \begin{bmatrix} -1 & 1 & & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ e^{i2\pi\mathbf{k}\cdot\mathbf{a}} & & & & -1 \end{bmatrix} \in \mathbb{C}^{m \times m}.$$

It is of course clear now that

- (i) The matrices $C_1, C_2,$ and C_3 are the discretizations of the operators ∂_x, ∂_y and ∂_z at the central face points, respectively.
- (ii) C_1^*, C_2^* and C_3^* are the discretizations of the operators $-\partial_x, -\partial_y$ and $-\partial_z,$ at the central edge points, respectively.
- (iii) The matrices C^*C and $I_3 \otimes (G^*G), G = [C_1^T \ C_2^T \ C_3^T]^T,$ are the discretizations of the operators $\nabla \times \nabla \times$ and $-\nabla^2$ at the central face points, respectively.

How do we effectively solve (12) using its eigenstructure is the focal point for the remaining part of this paper. We will first find eigendecomposition of the left hand side matrix. The fact that the singular value decomposition (SVD) of C can be obtained by Fast Fourier Transform (FFT) is the most important feature of our approach in solving problem (12). It is well known that the eigendecompositions of C^*C and CC^* are closely related to the SVD of C . The eigendecomposition of C^*C with the face-centered cubic (FCC) lattice translation vectors has been developed in [11]. We will use the techniques of [11] to derive the eigendecompositions of C^*C and CC^* for the simple cubic (SC) lattice with $C_\ell, \ell = 1, 2, 3$ in (13). The idea is to first ask if these C_ℓ can be diagonalized simultaneously. Recall from the earlier remark that these matrices are discrete versions of partial differentiation and so commutativity of multiplication among them should be expected, especially in the (symmetric) simple cubic case. We spell out the simultaneous diagonalization below. Recall that the exponentials are the eigenfunctions of partial operators and the Bloch's condition, the following quantities reflect these features in the discrete case. Define

$$(16) \quad \begin{aligned} \theta_{m,i} + \theta_{\mathbf{a},m} &\equiv \frac{i2\pi i}{m} + \frac{i2\pi \mathbf{k} \cdot \mathbf{a}}{m}, \\ D_{\mathbf{a},m} &= \text{diag} \left(1, e^{\theta_{\mathbf{a},m}}, \dots, e^{(m-1)\theta_{\mathbf{a},m}} \right), \\ \mathbf{u}_{m,i} &= [1 \quad e^{\theta_{m,i}} \quad \dots \quad e^{(m-1)\theta_{m,i}}]^\top \end{aligned}$$

for $i = 0, \dots, m-1$ and

$$(17) \quad \begin{aligned} U_m &= [\mathbf{u}_{m,0} \quad \dots \quad \mathbf{u}_{m,m-1}] \in \mathbb{C}^{m \times m}, \\ \Lambda_{\mathbf{a},m} &= \text{diag} \left(e^{\theta_{m,0} + \theta_{\mathbf{a},m}} - 1 \quad \dots \quad e^{\theta_{m,m-1} + \theta_{\mathbf{a},m}} - 1 \right). \end{aligned}$$

By the definition of $K_{\mathbf{a},m}$ in (15), it can be easily verified that

$$K_{\mathbf{a},m} (D_{\mathbf{a},m} U_m) = (D_{\mathbf{a},m} U_m) \Lambda_{\mathbf{a},m}.$$

Let

$$(18) \quad T = \frac{1}{\sqrt{n}} (D_{\mathbf{a}_3, n_3} \otimes D_{\mathbf{a}_2, n_2} \otimes D_{\mathbf{a}_1, n_1}) (U_{n_3} \otimes U_{n_2} \otimes U_{n_1}).$$

Note that the matrix-vector multiplication of T (by a vector) can be reduced to several one dimensional

cases by the following two formulas of Kronecker product

$$\begin{aligned} (A \otimes B \otimes C)(D \otimes E \otimes F) &= (AD) \otimes (BE) \otimes (CF), \\ (B^T \otimes A) \text{vec}(X) &= \text{vec}(AXB). \end{aligned}$$

where vec is the vectorization function. We remark in passing that if the periodic structure is a FCC lattice the matrix-vector multiplication of T can also be reduced to several one dimensional cases, the details can be found in [11]. Moreover, the complexity of matrix-vector multiplication of T can be accelerated from $O(n^2)$ to $O(n \log n)$ in SC and FCC by FFT.

Now the eigendecomposition of $C_\ell, \ell = 1, 2, 3$ can be stated as

Theorem 1 (Eigendecompositions of $C_\ell, \ell = 1, 2, 3$). *The unitary matrix T in (18) can simultaneously diagonalize $C_\ell, \ell = 1, 2, 3$:*

$$\begin{aligned} (19a) \quad C_1 T &= \delta_x^{-1} T (I_{n_3} \otimes I_{n_2} \otimes \Lambda_{\mathbf{a}_1, n_1}) \equiv T \Lambda_1, \\ (19b) \quad C_2 T &= \delta_y^{-1} T (I_{n_3} \otimes \Lambda_{\mathbf{a}_2, n_2} \otimes I_{n_1}) \equiv T \Lambda_2, \\ (19c) \quad C_3 T &= \delta_z^{-1} T (\Lambda_{\mathbf{a}_3, n_3} \otimes I_{n_2} \otimes I_{n_1}) \equiv T \Lambda_3, \end{aligned}$$

where $\Lambda_{\mathbf{a}_\ell, n_\ell}$ are from (17).

Thus, the null spaces of C^*C and CC^* can be derived immediately. To this end define

$$(20) \quad Q_0 = (I_3 \otimes T) \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \\ \Lambda_3 \end{bmatrix} \Lambda_q^{-1/2} \equiv (I_3 \otimes T) \Pi_0, \quad P_0 = (I_3 \otimes T) \overline{\Pi_0},$$

where $\Lambda_q = \Lambda_1^* \Lambda_1 + \Lambda_2^* \Lambda_2 + \Lambda_3^* \Lambda_3$. Then Q_0 and P_0 form orthogonal bases for the null spaces of C^*C and CC^* , respectively. Here we need an additional assumption so that Λ_q is positive definite: Assume $\mathbf{k} = (k_1, k_2, k_3)^\top \neq 0$ with $0 \leq k_1, k_2, k_3 \leq \frac{1}{2}$.

Next, we apply the techniques developed in [7, 11] to form the orthogonal bases for the range spaces of C^*C and CC^* . Firstly, letting $\Lambda_p = [\beta \Lambda_3 - \Lambda_2, \Lambda_1 - \alpha \Lambda_3, \alpha \Lambda_2 - \beta \Lambda_1]^\top$, considering the full column rank matrix $T_1 = [\alpha T^\top, \beta T^\top, T^\top]^\top$ with nonzero α and β , and taking the orthogonal projection of T_1 onto Q_0 and P_0 , respectively, we have

$$(21a) \quad \begin{aligned} Q_1 &= (I - Q_0 Q_0^*) T_1 (\Lambda_p^* \Lambda_p \Lambda_q^{-1})^{-1/2} \\ &= (I_3 \otimes T) \begin{bmatrix} (\alpha \Lambda_2 - \beta \Lambda_1) \Lambda_2^* - (\Lambda_1 - \alpha \Lambda_3) \Lambda_3^* \\ (\beta \Lambda_3 - \Lambda_2) \Lambda_3^* - (\alpha \Lambda_2 - \beta \Lambda_1) \Lambda_1^* \\ (\Lambda_1 - \alpha \Lambda_3) \Lambda_1^* - (\beta \Lambda_3 - \Lambda_2) \Lambda_2^* \end{bmatrix} \\ &\quad \times (\Lambda_p^* \Lambda_p \Lambda_q^{-1})^{-1/2} \\ &\equiv (I_3 \otimes T) \Pi_1, \end{aligned}$$

$$(21b) \quad P_1 = (I - P_0 P_0^*) T_1 (\Lambda_p^* \Lambda_p \Lambda_q^{-1})^{-1/2} = (I_3 \otimes T) \overline{\Pi_1}.$$

Then Q_1 and P_1 are orthogonal, and $(C^*C)Q_1 = Q_1 \Lambda_q$ and $(CC^*)P_1 = P_1 \Lambda_q$. Secondly, in order to form the remaining part of the orthogonal basis for the range

spaces of C^*C and CC^* , we apply the discrete dual-curl and curl operators on T_1 , respectively. That is, we premultiply T_1 with C^* and C , respectively, to get

$$(22a) \quad Q_2 = C^*T_1 (\Lambda_p^* \Lambda_p)^{-1/2} \\ = (I_3 \otimes T) \begin{bmatrix} \beta \Lambda_3^* - \Lambda_2^* \\ \Lambda_1^* - \alpha \Lambda_3^* \\ \alpha \Lambda_2^* - \beta \Lambda_1^* \end{bmatrix} (\Lambda_p^* \Lambda_p)^{-1/2} \\ \equiv (I_3 \otimes T) \Pi_2, \\ (22b) \quad P_2 = CT_1 (\Lambda_p^* \Lambda_p)^{-1/2} = (I_3 \otimes T) (-\overline{\Pi_2}).$$

One can easily check that $(C^*C)Q_2 = Q_2\Lambda_q$ and $(CC^*)P_2 = P_2\Lambda_q$. From (20)-(22), we define

$$(23a) \quad Q \equiv [Q_1 \quad Q_2 \quad Q_0] = (I_3 \otimes T) [\Pi_1 \quad \Pi_2 \quad \Pi_0], \\ (23b) \quad P \equiv [P_2 \quad P_1 \quad P_0] = (I_3 \otimes T) [-\overline{\Pi_2} \quad \overline{\Pi_1} \quad \overline{\Pi_0}].$$

Finally, the eigendecomposition of C^*C can be formally stated as follows.

Theorem 2 (Eigendecompositions of C^*C). *Let Λ_i , $i = 1, 2, 3$ be defined as in (19a)-(19c). Then*

$$(24) \quad C^*C = Q \text{diag}(\Lambda_q, \Lambda_q, 0) Q^*$$

where the unitary Q is from (23) and $\Lambda_q = \Lambda_1^* \Lambda_1 + \Lambda_2^* \Lambda_2 + \Lambda_3^* \Lambda_3$.

A similar statement holds for CC^* as well with Q replaced by P . From this statement and (24), we can construct the left and right singular vector matrices P and Q for C with an optimal complexity if possible. Note that since nonzero eigenvalues of C^*C have even multiplicities that can be very large, the choices of bases for P and Q in (23) are not unique. The special choices of bases for P and Q in (23) are nontrivial and essential for the practical availability of the SVD of C .

Theorem 3 (Singular value decomposition of C). *Let Λ_q and (Q, P) be defined in (23). Assume that the vector $\mathbf{k} = (k_1, k_2, k_3)^T$ is nonzero with $0 \leq k_1, k_2, k_3 \leq \frac{1}{2}$. Then, the matrix C has the SVD*

$$(25) \quad C = P \text{diag}(\Lambda_q^{1/2}, \Lambda_q^{1/2}, 0) Q^* = P_r \Sigma_r Q_r^*,$$

where

$$P_r = [P_2, P_1], \quad Q_r = [Q_1, Q_2], \quad \Sigma_r = \text{diag}(\Lambda_q^{1/2}, \Lambda_q^{1/2}),$$

where $P_i, Q_i, i = 1, 2$ are from (23).

We now give an example to show how the eigendecomposition of C^*C can be used to solve GEP (12). Let us assume that the material under consideration is non-magnetic ($D_\mu = I$) and is such that B_ζ and B_ξ in (12) are the zero matrix. Then problem (12) becomes a discrete double curl GEP for the electric field \mathbf{e} :

$$(26) \quad C^*C\mathbf{e} = \omega^2 D_\epsilon \mathbf{e},$$

which has a large null space by Theorem 2. The same theorem suggests a powerful technique to reduce (26) to a null space free eigenvalue problem (NFEP).

A general transformation from the GEP to the NFEP has been shown in [11]. In the present context it is based on the following theorem.

Theorem 4 (GEP to NFEP). *Let*

$$A = C^*C, \quad B = D_\epsilon, \quad \text{and} \quad \Lambda_r = \text{diag}(\Lambda_q, \Lambda_q).$$

Then we have

$$(27) \quad \text{span} \left\{ B^{-1} Q_r \Lambda_r^{\frac{1}{2}} \right\} = \{x \mid Ax = \lambda Bx, \lambda > 0\}$$

and

$$(28) \quad \{\lambda \mid Ax = \lambda Bx, \lambda > 0\} = \{\lambda \mid A_r y = \lambda y\},$$

where

$$A_r = \Lambda_r^{\frac{1}{2}} Q_r^* B^{-1} Q_r \Lambda_r^{\frac{1}{2}}.$$

In fact, from Theorem 2 and notation in Theorem 3, we have $A = Q_r \Lambda_r Q_r^*$. Now

$$(B^{-1}A) \left(B^{-1} Q_r \Lambda_r^{1/2} \right) = (B^{-1} Q_r \Lambda_r Q_r^*) \left(B^{-1} Q_r \Lambda_r^{1/2} \right) \\ = \left(B^{-1} Q_r \Lambda_r^{1/2} \right) A_r,$$

and consequently $\text{span}\{B^{-1}Q_r\Lambda_r^{1/2}\}$ is an invariant subspace of $B^{-1}A$ associated with the eigenvalues of A_r . Since A_r is positive definite and has rank $2n$, assertions (27) and (28) hold.

Now applying the last theorem to (26) we arrive at

$$(29) \quad A_r \mathbf{u} = \omega^2 \mathbf{u}$$

where

$$A_r = \Lambda_r^{\frac{1}{2}} Q_r^* D_\epsilon^{-1} Q_r \Lambda_r^{\frac{1}{2}} \quad \text{and} \quad \mathbf{u} = \Lambda_r^{-\frac{1}{2}} Q_r^* D_\epsilon \mathbf{e}.$$

This technique is called null space free method [7, 11], which is capable of reducing (26) to a problem like (29) that contains only all the nonzero invariant eigen-subspaces. Due to the facts that the matrix-vector multiplication of Q_r can be accelerated by FFT, and that the matrix A_r is positive definite with small condition numbers (two digits sometimes), problem (29) is much easier to solve than the original problem (26). Furthermore because the special choices of the bases mentioned before, this method requires only the storage of diagonal matrices and vectors, which significantly save the computer storage space.

If the material is metal, the permittivity ϵ should be considered as a function of frequency ω , and one

of the most often used model of permittivity for a metallic material is the Drude model [10, 17] in which

$$(30) \quad \varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\omega\gamma},$$

where ω_p is the plasma frequency and γ is the collision frequency or damping constant. Note that with γ small ε could be “negative”. Under this assumption, (12) can be transformed into a polynomial eigenvalue problem for which Jacobi-Davidson method or Residual Arnoldi method is applicable. No matter what method one adopts, there are some linear systems that must be solved along the way. Fortunately, the technique of FFT can be used to design good preconditioners so that these linear systems can be efficiently solved. Further details are discussed in [7].

In addition, if ζ and ξ are mutually complex conjugate in problem (12), then the material is a special case of complex media called Reciprocal Chiral medium or Chiral medium. Further if $D_\varepsilon - B_\xi D_\mu^{-1} B_\zeta > 0$ (Hermitian positive definite), then by using the null space free method to remove the null space of (12), problem (12) can also be simplified into the following form

$$(31) \quad \iota \begin{bmatrix} 0 & -\Sigma_r^{-1} \\ \Sigma_r^{-1} & 0 \end{bmatrix} y_r = \frac{1}{\omega} (R^* R) y_r,$$

where matrix R is the range space of (12) and can be written as

$$R = \begin{bmatrix} (D_\varepsilon - B_\xi D_\mu^{-1} B_\zeta)^{-\frac{1}{2}} & 0 \\ 0 & D_\mu^{-\frac{1}{2}} \end{bmatrix} \times \begin{bmatrix} B_\xi D_\mu^{-1} & I_{3n} \\ -I_{3n} & 0 \end{bmatrix} \text{diag}(P_r, Q_r).$$

Since $\iota \begin{bmatrix} 0 & -\Sigma_r^{-1} \\ \Sigma_r^{-1} & 0 \end{bmatrix}$ is symmetric and since $R^* R$ is positive definite, the matrix-vector multiplication of matrix R still can be accelerated by FFT, so that the computational cost remains low and the storage can also be mostly reduced. Once again all these advantages are due to the FFT and the singular value decomposition of the curl matrix C . Further details of such applications can be found in [7].

Discrete Maxwell Eigenvalue Problem by Edge Elements

The finite element method has been used extensively in engineering and science, but its use in the fundamental research of photonic crystals is more recent [25] and a critical assessment of the finite element method for two dimensional photonic crystals can be found in [1]. The mathematical theory of finite element methods for Maxwell equations is subtle and can be found in [14]. We will now directly derive a dis-

crete weak formulation from a computational viewpoint, avoiding using the language of $H(\text{curl})$ spaces and trace operators via duality, which is necessary if a continuous weak formulation were to be found first. The main point of the informal approach is to bring out the interplay between interelement boundary conditions and between outerface boundary conditions and Bloch’s periodic boundary conditions.

Consider a nonmagnetic material with the constitutive relation (3), so that (6) is valid over a domain Ω in \mathbb{R}^n :

$$(32) \quad \nabla \times \nabla \times E = \lambda \varepsilon E,$$

appended with a divergence free condition. Note that if $\lambda \neq 0$ then the divergence free condition $\nabla \cdot \varepsilon E = 0$ is naturally satisfied. On the other hand if $\lambda = 0$, then those E generated by potentials i.e., $E = \nabla \phi$ are in the null space of the double curl operator. Whether the space is large or trivial hinges on the boundary condition associated with the Poisson equation $\nabla \cdot \varepsilon \nabla \phi = 0$. When Bloch’s quasi-periodic condition is enforced, a large null space arises as a result. The null space is trivial when on the boundary ϕ is constant or equivalently the tangential component of E is zero. A good discrete weak formulation of (32) will respect these features as well as discontinuity in the material parameter ε due to the presence of different dielectric or metallic materials in a photonic crystal.

Let

$$\bar{\Omega} = \cup_{r=1}^n K_r,$$

be a subdivision of Ω into a union of simplices and let

$$(f, g)_K = \int_K f(\mathbf{x}) \bar{g}(\mathbf{x}) d\mathbf{x}, \quad (f, g)_{\partial K} = \int_{\partial K} f(\mathbf{x}) \bar{g}(\mathbf{x}) d\sigma$$

be the usual L^2 scalar products on a region K and its boundary ∂K , respectively. Using vector identities

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \quad \text{and} \quad \nabla \cdot (\mathbf{a} \times \mathbf{b}) = (\nabla \times \mathbf{a}) \cdot \mathbf{b} - \mathbf{a} \cdot (\nabla \times \mathbf{b}),$$

we can deduce the integration-by-parts formula

$$(33) \quad \begin{aligned} (\nabla \times \nabla \times E, W)_K &= (\nabla \times E, \nabla \times W)_K + (\mathbf{n} \times \nabla \times E, W)_{\partial K} \\ &= (\nabla \times E, \nabla \times W)_K + (\mathbf{n} \times \nabla \times E, W_T)_{\partial K}, \end{aligned}$$

where

$$W_T = (\mathbf{n} \times W) \times \mathbf{n}$$

is the tangential projection of W to ∂K and \mathbf{n} is the unit exterior normal on ∂K . Here we have assumed E and W have the the required smoothness. Now applying (33) with $K = K_r$ for all r and summing, we arrive at a weak formulation

$$(34) \quad (\nabla \times E, \nabla \times W)_\Omega + \sum_r (\mathbf{n} \times \nabla \times E, W_T|_{K_r})_{\partial K_r} = \lambda \varepsilon (E, W)_\Omega,$$

where $W_T|_{K_r}$ is W_T from the side of K_r . The second term on the left contains contributions of surface integrals from interior surfaces common to two K_r 's and surface integrals from the boundary of Ω . The interior portions can be made zero pairwise due to the opposite signs of the two normals, provided that we choose W to have continuous tangential components across the interfaces and assume E is smooth. The boundary portions can be handled by the imposed boundary condition. On the one hand, if the whole boundary is also the boundary of a unit cell of a lattice then the boundary condition is a natural boundary condition and Bloch's condition should hold for W if we adopt Galerkin method in which the test space is the same as the trial space, the electric field space. On the other hand, if the unit cell is just a part of the computational domain Ω , then part of the boundary may contain Dirichlet condition for $E \times \mathbf{n}$. Of course, there are other possible boundary conditions that may be imposed. With this in mind, the finite element test space is taken to be the edge elements of Nedelec or Whitney [15, 6]. The discrete eigenvalue problem takes the form of finding $E_h \in \mathcal{N}_h$ such that

$$(35) \quad a(E_h, W) = \lambda \varepsilon(E_h, W)_\Omega, \quad \forall W \in \mathcal{N}_h,$$

where

$$(36) \quad a(E_h, W) = (\nabla \times E_h, \nabla \times W)_\Omega + (\mathbf{n} \times \nabla \times E_h, W_T)_{\partial\Omega}.$$

In terms of basis functions $\{W_i\}$, we have a GEP

$$A\mathbf{x} = \lambda B\mathbf{x},$$

where $A_{ij} = a(W_i, W_j)$ and $B_{ij} = \varepsilon(W_i, W_j)_\Omega$. If the boundary term on the right of (36) is zero (e.g., PCE boundary condition imposed) then both A and B are Hermitian and the GEP can be directly solved by Arnoldi's method and some basic linear solvers. Does the SVD approach in the previous section carry over? Edge elements on bricks (hexahedra) are equivalent to the Yee cells as hinted by notations following (9) (cf. [5, 9]). Extension to edge elements on tetrahedrons is underway. Now let us look at the case when (35) is to be solved on Ω , a unit cell of an ideal SC lattice. Thus no degrees of freedom are to be imposed on $\partial\Omega$. Using brick edge elements one can check that the boundary term on the right of (36) is zero, since $\mathbf{n} \times \nabla \times W_i = 0$ on boundary surfaces. Hence the matrix A is Hermitian. For the unknowns \mathbf{x} , the boundary unknowns are related by Bloch's condition:

$$\mathbf{x}_+ = e^{2\pi i \mathbf{k} \cdot \mathbf{a}} \mathbf{x}_-,$$

where $(+, -)$ can be the top-bottom, right-left, or front-rear pair. Setting $\tilde{\mathbf{x}} = (\mathbf{x}_0, \mathbf{x}_l, \mathbf{x}_b, \mathbf{x}_r)^\top$ into the interior part and boundary parts and expressing the ex-

ponential multiplication as a matrix, we turn the Hermitian system $A\mathbf{x} = \lambda B\mathbf{x}$ into another Hermitian one:

$$P^* A P \tilde{\mathbf{x}} = \lambda P^* B P \tilde{\mathbf{x}}.$$

Extension of the SVD approach to this form for brick and tetrahedral edge elements is underway.

A final remark is in order here. We can always start with a weak formulation of the continuous model using integration by parts and arrive at

$$(\nabla \times E, \nabla \times W)_\Omega + (\mathbf{n} \times \nabla \times E, W_T)_{\partial\Omega} = \lambda \varepsilon(E, W)_\Omega.$$

The boundary term $(\mathbf{n} \times \nabla \times E, W_T)_{\partial\Omega}$ can be further simplified, depending on the boundary conditions imposed. Actually we can drop the contribution from the pair (top, bottom), (left, right) or (front, rear) surfaces if on that pair the Bloch condition is imposed. For example, if the condition is on the (R = right, L = left) pair, then

$$\begin{aligned} (\mathbf{n}_R \times \nabla \times E|_R, W_T|_R) &= (-\mathbf{n}_L \times \nabla \times e^{2\pi i \mathbf{k} \cdot \mathbf{a}} E|_L, e^{2\pi i \mathbf{k} \cdot \mathbf{a}} W_T|_L) \\ &= -(\mathbf{n}_L \times \nabla \times E|_L, W_T|_L). \end{aligned}$$

Thus

$$(\mathbf{n} \times \nabla \times E, W_T)_{R \cup L} = 0.$$

If the Bloch condition is imposed on $\partial\Omega$ and a conforming approximation such as the edge element is used, we then have a Hermitian system.

Numerical Results

In this section, we present some numerical results of band structures for perfect vacuum state, photonic crystal and chiral medium with simple cubic lattice structure. Yee's scheme is used in the discretization. The medium is composed of cylinders and spheres. In this case, the first Brillouin zone of the simple cubic lattice is formed by the corners

$$\begin{aligned} G &= [0 \ 0 \ 0]^T, \quad X = \frac{2\pi}{a} [\frac{1}{2} \ 0 \ 0]^T, \\ M &= \frac{2\pi}{a} [\frac{1}{2} \ \frac{1}{2} \ 0]^T, \quad R = \frac{2\pi}{a} [\frac{1}{2} \ \frac{1}{2} \ \frac{1}{2}]^T, \end{aligned}$$

where a is the side length of the simple cubic lattice. The permittivity of material is ε_i and outside of the material is air with permittivity ε_0 . We set the matrix size in (12) to 4×50^3 . We first show the band structure with constitutive relations (2).

Since in this case the permittivity is ε_0 everywhere, in Figure 1 we see that there does not exist any band gap in low frequencies, which is both intuitively and theoretically expected. So we pay attention to the next case in which the material is made up of spheres that are connected by circular cylinders.

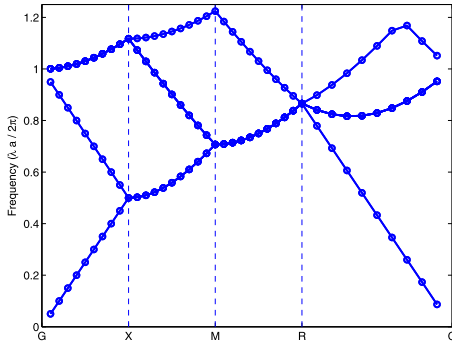


Figure 1. The band structure with $\varepsilon_i = 1, \zeta = \xi = 0$.

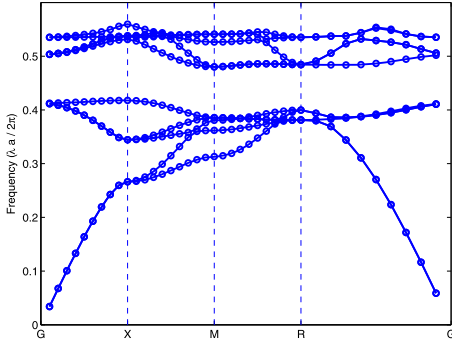


Figure 2. The band structure with $\varepsilon_i = 13, \zeta = \xi = 0$.

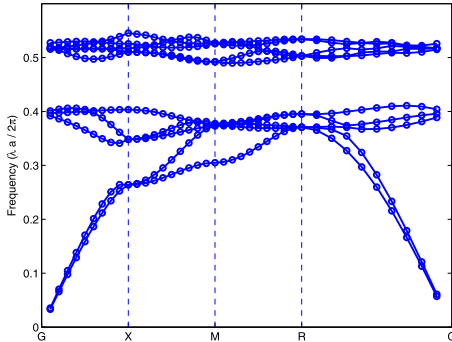


Figure 3. The band structure with $\varepsilon_i = 13, \zeta = 0.5i, \xi = -0.5i$.

The associated permittivity ε_i is set to 13, and ζ and ξ set to zero. We can see in Figure 2 that there appears a complete band gap around $\omega = 0.4$ to 0.5 . In Figure 3, we see what happens if the material is chosen to be a complex medium. In this figure, the associated permittivity ε_i is still set to 13. In addition, ζ and ξ are set to $0.5i$ and $-0.5i$, respectively.

We can observe that each eigen-curve splits into several eigen-curves and that the band structure seems to be wider than the band structure in Figure 2. However, the above figures are just some preliminary examples, we believe that there are more interesting

physical phenomena to be found by expanding and improving the present approaches.

Conclusions

We have reported some recent advances [7, 11] and their extensions in the computation of the band structures of photonic crystals. Our approach is to deduce explicit eigendecompositions for the discrete single and double curl operators associated with the eigenvalue problems resulting from the discretization of the Maxwell eigenvalue problem. One major strength of our method is that these decompositions can be efficiently computed by FFT; resulting in novel creation of fast eigensolvers. Another strength is that the efficient computational realization of the null space enables us to derive the null space free method to avoid the commonly annoying issue of large null space, when it comes to applying iterative eigenvalue solvers. Furthermore, the reduction of the problem size is often substantial. All these claims are further supported by the numerical results.

Overall, the strengths of our approach can be attributed to the success and better understanding of the following two steps:

- explicit constructions of eigenstructures of the discrete single and double curl operators.
- judicious design of bases for the associated eigenspaces that makes application of the FFT possible.

Our future works will concentrate on further developments and generalizations of these two steps in underlying problems. Step *a* is independent of the application of photonic crystals. So far our discrete curl operators are all based on either Yee's finite difference scheme or its counterpart in the edge element finite element schemes. While these schemes have shown to be very robust, they are only of lower order accuracy. When smoother solutions are involved it is important to investigate the issue for higher order edge elements. Also since the discrete curl operators not only depend on the discretization but also on the underlying boundary conditions, we want to generalize this step to include Dirichlet, Neumann, and PML conditions associated with not only eigenvalue problems but also source problems. Explicit construction and deeper understanding of the eigenstructure for the curl-type discrete operator resulting from the idea of complex stretching in the PML case would constitute a major advance in the related fields. All these issues are currently under investigation in our project outlined in [8]. Concerning Step *b*, we note that the fact that the computational domain is a periodic structure plays a central role and how to accelerate computation of the matrix-vector multipli-

cation via FFT often depends on a good understanding of the tensor representations involved. Finally, we point out that while there are other existing methods for Maxwell equations [4] such as the (discontinuous) Galerkin method and mixed methods, none have considered the above two steps together in their design. Without considering them together, it seems much harder to design fast solvers.

Acknowledgements

This work is partially supported by the National Science Council, the National Center for Theoretical Sciences and the Chiao-Da ST Yau Center in Taiwan.

References

- [1] I. Andonegui and A. J. Garcia-Adeva. The finite element method applied to the study of two-dimensional photonic crystals and resonant cavities. *Optics Express*, 21(4):4072, 2013.
- [2] J. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185-200, 1994.
- [3] F. Bloch. Über die quantenmechanik der elektronen in kristallgittern. *Z. Physik*, 52:555-600, 1928.
- [4] D. Boffi. Finite element approximation of eigenvalue problems. *Acta Numerica*, 19:1-120, 2001.
- [5] A. Bonderson, T. Rylander, and P. Ingelstrom. *Computational electromagnetics*. Springer, 2005.
- [6] A. Bossavit. *Mixed finite elements and the complex of Whitney forms in The Mathematics of finite elements and applications VI*. Academic Press, San Diego, CA, 1998.
- [7] R. L. Chern, H. E. Hsieh, T. M. Huang, W. W. Lin, and W. Wang. Singular value decompositions for single-curl operators in three-dimensional Maxwell's equations for complex media. *SIAM Matrix Anal. Appl.*, 36(1): 203-224, 2015.
- [8] S. H. Chou, H. E. Hsieh, and W. W. Lin. Eigendecompositions of discrete curl-type operators and their realization by FFT. Technical report, National Center for Theoretical Sciences, National Tsing Hua University, Taiwan, Preprints in Mathematics 2014-6-001, 2014.
- [9] G. C. Cohen. *Higher-Order numerical methods for transient wave equations*. Springer, 2002.
- [10] P. Drude. Zur elektronentheorie der metalle. *Annalen der Physik*, 306(3):566, 1900.
- [11] T. M. Huang, H. E. Hsieh, W. W. Lin, and W. Wang. Eigendecomposition of the discrete double-curl operator with application to fast eigensolver for three dimensional photonic crystals. *SIAM Journal on Matrix Analysis and Applications*, 34(2):369-391, 2013.
- [12] C. Kittel. *Introduction to solid state physics*. Wiley, New York, 2005.
- [13] S. Maier. *Plasmonics: fundamentals and applications*. Springer, 2007.
- [14] P. Monk. *Finite element methods for Maxwell's equations*. Clarendon Press, Oxford, 2003.
- [15] J. C. Nedelec. Mixed finite elements in \mathbb{R}^3 . *Numerische Mathematik*, 35(3):315-341, 1980.
- [16] M. A. Noginov and V. A. Podolskiy. *Tutorials in metamaterials*. CRC Press, 2012.
- [17] J. B. Pendry, A. J. Holden, W. J. Stewart, and J. Young. Extremely-low-frequency plasmons in metallic mesostructures. *Phys. Rev. Lett.*, 306(76):4773, 1996.
- [18] R. A. Shelby, D. R. Smith, and S. Schultz. Experimental verification of a negative index of refraction. *Science*, 292(5514):77-79, 2001.
- [19] D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz. Composite medium with simultaneously negative permeability and permittivity. *Physical Review Letters*, 84(18):4184-4187, 2000.
- [20] V. G. Veselago. The electrodynamics of substances with simultaneously negative values of ϵ and μ . *Sov. Phys. Usp.*, 10(4):509-14, 1968 (Russian text 1967).
- [21] J. Volakis, A. Chatterjee, and L. C. Kempel. *Finite element method for electromagnetics: antennas, microwave circuits, and scattering applications*. New York: Wiley IEEE, 1998.
- [22] E. Yablonovitch. Inhibited spontaneous emission in solid-state physics and electronics. *Physical Review Letters*, 58(20):2059-2062, 1987.
- [23] E. Yablonovitch, T. J. Gmitter, and K. M. Leung. Photonic band structure: the face-centered-cubic case employing nonspherical atoms. *Physical Review Letters*, 67(17):2295-2298, 1991.
- [24] K. 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.
- [25] F. Zolla, G. Renversez, A. Nioclet, B. Kuhlmeiy, S. Guenneau, and D. Felbacq. *Foundtions of photonic crystal fibers*. World Scientific, 2005.