Topology optimization of 3D photonic crystals with complete bandgaps

KENNETH E. SWARTZ,[1,2,*] ⃝ DANIEL A. WHITE,[1] DANIEL A. TORTORELLI,[1,2] AND KAI A. JAMES[3]

[1]*Lawrence Livermore National Laboratory, Computational Engineering Division, 7000 East Ave, Livermore, CA 94550, USA*

[2]*University of Illinois Urbana-Champaign, Department of Mechanical Science and Engineering, 1206 W Green St, Urbana, IL 61801, USA*

[3]*University of Illinois Urbana-Champaign, Department of Aerospace Engineering, 104 S Wright St, Urbana, IL 61801, USA*

*swartz10@llnl.gov

Abstract: The design of photonic crystals with complete bandgaps has recently received considerable research focus for numerous reasons. This work leverages well-known nonlinear programming techniques to alleviate the non-smoothness caused by degenerate eigenvalues such that topology optimization problems can be solved with the open-source IPOPT software. A fully-vectorial plane wave expansion technique is used with an iterative eigensolver to efficiently predict dispersion properties of candidate structures. Nonlinear programming is employed to solve the inverse problem of designing three-dimensional periodic structures that exhibit complete two-dimensional (2D) and three-dimensional (3D) photonic bandgaps. Mesh refinement is performed to alleviate the large computational burden of designing and analyzing photonic crystals, and a periodic density filter is implemented to impose a minimum feature size for manufacturability considerations.

1. Introduction

Photonic bandgap structures forbid electromagnetic wave propagation of a given frequency range, also known as the bandgap [1]. The first photonic bandgap structures were one-dimensional periodic structures [2], however, the bandgap engineering field progressed little until higher-dimensional bandgap structures were discovered a century later [3,4]. Early experiments demonstrated 3D photonic bandgaps in crystals made of dielectric spheres with face-centered cubic [5] and body-centered cubic [6] symmetries, as well as in crystals made of dielectric rods [7]. Additionally, a diamond structure of dielectric spheres was designed and its complete 3D photonic bandgap was numerically simulated [8]. Since these groundbreaking results, three decades of research has followed on the design and analysis of photonic bandgap structures. Novel applications include low-loss waveguides [9–12], perfect mirrors [13], photonic cavities [11,14], photonic bandgap fibers [15], and a number of other exciting optical applications [1].

A material made of a periodic, architected micro-structure is referred to as a "metamaterial". The bulk, or effective, material properties of metamaterials are controlled by the underlying material properties, but also and more interestingly, by the geometry of the microstructure. A technique known as homogenization was developed to predict the effective elastic properties of metamaterials by analyzing just a single unit cell [16]. Further, topology optimization [17,18], which seeks to determine the optimal material distribution for a specific task, is readily applied to solve the inverse homogenization problem [19,20], i.e. to generate micro-structures that exhibit optimal, or prescribed, effective properties. Seemingly impossible effective material properties have been observed by optimizing metamaterials, such as simultaneous negative bulk modulus and mass density [21] as well as negative Poisson's ratio [22].

Metamaterials have also been designed for effective electromagnetic properties, such as magnetic permeability and electric permittivity, although much less research has been performed in this realm. So called "left-handed" metamaterials have been designed to exhibit negative effective permeability and permittivity [23]. These electromagnetic metamaterials are periodic structures that exhibit their effective bulk properties at an operating frequency with a corresponding wavelength that is much larger than the unit cell size [24]. Photonic crystals are similarly periodic structures, however, they obtain their performance by leveraging diffraction [24] and are therefore designed to operate at frequencies with a wavelength that is similar in size to the unit cell. Numerical inverse design, referred to as inverse homogenization via topology optimization, is readily applied to design the periodic unit cells of both metamaterials and photonic crystals.

Many previous studies have applied inverse problem techniques to the design of photonic crystals [25]. A major issue plaguing numerical design of photonic bandgap structures is the non-smoothness of the objective function due to degenerate eigenvalues [26]. The first successful 3D bandgap design optimization naturally avoided eigenvalue degeneracy by using a very low-dimensional design space [27], however, a number of techniques have since been employed to circumvent this issue. 2D photonic crystals exhibiting bandgaps were designed using a generalized gradient ascent [28,29] algorithm that leveraged directional derivatives, or subgradients, computed by the generalized gradient technique for degenerate eigenvalues [30]. The generalized gradient ascent was later extended to a level-set design parameterization [31], and similarly a regularized descent method was used to design optical waveguides via shape optimization [32]. Alternatively, semidefinite programming, which alleviates the need for eigenvalue sensitivities, was applied to design 2D photonic crystals [33] and further extended to the robust design of 3D photonic crystals to accommodate uncertainties in fabrication [34]. Topology optimization, i.e. inverse unit cell design, has also been used to design structures with mechanical, or phononic, bandgaps [35], and even more interestingly with both phononic *and* photonic bandgaps [36]. Genetic [37] and evolutionary [38] topology optimization methods have also been applied to design photonic crystals. These techniques do not require design sensitivities, but do require many more iterations to converge than gradient based-algorithms and they cannot ensure optimality. For these reasons, gradient-free algorithms are not recommended for solving topology optimization problems [39]. At the other extreme, theoretical techniques for bandgap design that do not require iterative techniques were developed [40,41]. Interested readers are directed to the review articles [42,43] for further exposition.

The numerical techniques described in this work are only applicable to non-magnetic materials, and therefore we direct readers to [44–48] for studies on magnetic photonic crystals. Design optimization has not yet been applied to 3D magnetic photonic crystals. Further, we assume lossless materials and direct readers to in [1,49,50] for information on lossy photonic structures.

This work alleviates the differentiability issues of repeated eigenvalues by leveraging research on symmetry polynomials of eigenvalues [51]. The presented smoothing technique allows photonic crystal design via traditional gradient-based nonlinear programming (NLP) solvers. The use of a smooth NLP algorithm is advantageous, because although nonsmooth NLP algorithms [52] could be used in principle, their success has not been demonstrated on photonic crystal design in practice. Our work also uses a 3D design parameterization, which is necessary to generate structures with complete 3D bandgaps; it also proved very effective for optimizing structures with large complete 2D bandgaps. We also improved the computational efficiency of a "voxel" based design parameterization by using a nested mesh refinement technique. Length-scale control is implemented to ensure manufacturability by employing a periodic volume fraction filter that respects the orthorhombic symmetry of our unit cell. Multi-level parallelism and an efficient sensitivity analysis are exploited to execute these large computational tasks in a reasonable wall clock time. We validate our framework by designing novel structures with complete 2D and 3D bandgaps.

2. Solution methods

2.1. Maxwell's eigenvalue problem

Time-harmonic electromagnetism in a source-free, isotropic, lossless medium is governed by the four Maxwell equations [1,53]

$$\nabla \times E(x) = i\omega\mu_0\mu_r(x)H(x) \tag{1}$$

$$\nabla \times H(x) = -i\omega\varepsilon_0\varepsilon_r(x)E(x) \tag{2}$$

$$\nabla \cdot (\varepsilon_r(x)E(x)) = 0 \tag{3}$$

$$\nabla \cdot (\mu_r(x)H(x)) = 0, \tag{4}$$

where $\nabla\times$ and $\nabla\cdot$ are the curl and divergence operators with respect to position x, respectively, E is the electric field, H is the magnetic field, ω is the frequency, μ_0 is the magnetic permeability in a vacuum, μ_r is the relative magnetic permeability, ε_0 is the electric permittivity in a vacuum, ε_r is the relative electric permittivity, and i is the imaginary unit. Bold-faced font will be used to denote vector entities throughout. This paper will focus on non-magnetic materials, i.e. $\mu_r(x) = 1$, and therefore we omit $\mu_r(x)$ hereafter. Equations (1) and 2 are easily combined into

$$\nabla \times \left(\frac{1}{\varepsilon_r(x)}\nabla \times H(x)\right) = \left(\frac{\omega}{c}\right)^2 H(x), \tag{5}$$

where it is noted that the speed of light $c = \frac{1}{\mu_0\varepsilon_0}$ was substituted, and we reiterate that our media are assumed to be lossless and therefore ε_r is strictly real. Further, since we are assuming isotropic media we express the permittivity as a scalar quantity, although the solution technique easily generalizes to anisotropoic media wherein the permittivity is tensor-valued and thus $\frac{1}{\varepsilon_r(x)}$ becomes ε_r^{-1}. Equation (5) is an ordinary eigenvalue problem with eigenpairs $\left(\left(\frac{\omega}{c}\right)^2, H\right)$, however, it is important to remember that solutions to Eq. (5) must also satisfy Eqs. (3) and (4) to be physically valid. A similar eigenvalue problem for the electric field E could have been formulated, although a generalized eigenvalue problem would have been obtained. The choice of solution field, i.e. H or E, is further discussed in Sec. 2.2

Bloch's theorem [54,55] allows the magnetic field in a periodic medium to be expressed over a single unit cell as

$$H(x) = H_p(x)e^{i\kappa \cdot x}, \tag{6}$$

where H_p is a periodic function with the same periodicity as the medium and κ is a wave vector. Substituting Eq. (6) into Eq. (5) yields the ordinary eigenvalue problem

$$A_\kappa H_p(x) = (\nabla + i\kappa) \times \frac{1}{\varepsilon_r(x)}(\nabla + i\kappa) \times H_p(x) = \left(\frac{\omega}{c}\right)^2 H_p(x), \tag{7}$$

which is solved over a single unit cell for the pair $\left(\left(\frac{\omega}{c}\right)^2, H_p\right)$. From Eqs. (6) and (7) we see that H is a wave that propagates with frequency ω in the direction κ. In bandgap structures there exists intervals $[\omega_1, \omega_2]$, i.e. bandgaps, in which no waves propagate.

2.2. Planewave expansion

A finite dimensional basis is used to approximate H and solve Eq. (7) numerically. There are a number of possible bases, each with their own advantages and limitations [56]. Perhaps the most common basis in numerical methods is a finite element basis [57], which has the advantage of easily enforcing the required solution continuity through the use of vector finite elements, e.g. H-curl finite elements enforce tangential vector continuity [58]. Due to their immense applicability

and popularity across many disciplines, open source finite element codes are widely available, such as the moduluar finite element methods (MFEM) library [59]. The main disadvantage of a finite element basis is the difficulty of enforcing the transversality requirements [56]. As mentioned in Sec. 2.1, all computed solutions of Eq. (7) must also satisfy Eqs. (3) and (4) to be physically valid. Solving Eq. (7) numerically without a transversality constraint will yield many zero-valued eigenvalues [56] which do not satisfy Eqs. (3) and (4), These spurious solutions could be removed after the calculation, but this technique would be quite computationally wasteful. A much better choice is the use of a planewave basis, which has become the standard practice for photonic band analyses since it was first presented two decades ago [56].

A planewave basis is used to approximate H_p as

$$H_\mathrm{p}(x) = \sum_{m=1}^{N} h_m e^{iG_m \cdot x} \tag{8}$$

where G_m are the reciprocal lattice vectors of the unit cell and h_m are the unknown degrees of freedom (DOF) in wavevector space [56]. Substituting Eq. (8) into Eq. (6) and inspecting the resulting Eq. (4) reveals a major advantage of the planewave basis. We see that the transversality requirement reduces to

$$\nabla \cdot \left(\sum_{m=1}^{N} h_m e^{i(G_m + \kappa) \cdot x} \right) = 0, \tag{9}$$

which further reduces to

$$h_m \cdot (G_m + \kappa) = 0. \tag{10}$$

Since $G_m + \kappa$ is known, the above Eq. (10) constraint on h_m is easily enforced by expressing

$$h_m = h_m^1 e_1 + h_m^2 e_2 \tag{11}$$

where e_1 and e_2 are orthogonal unit vectors that are perpendicular to $G_m + \kappa$, and solving for the unknown DOF h_m^1 and h_m^2. The transversality constraint would not be so easily enforced had Eq. (5) been formulated for the electric field E, due to the presence of $\varepsilon_\mathrm{r}(x)$ in Eq. (3). By choosing H as the solution field, satisfying $\nabla \cdot H(x) = 0$ effectively satisfies Eq. (4), since $\mu_\mathrm{r}(x) = 1$, and Eq. (3) will automatically be satisfied by expressing h_m via Eq. (11) [56].

Bandgap optimizations only require calculation of a few of the lowest modes, and thus for large problems an interative eigenvalue solver is significantly more efficient than a direct solver. Fortunately, the linear operator in Eq. (7) is symmetric positive-definite [1], and as such iterative methods, e.g. locally optimal block preconditioned conjugate gradient (LOBPCG) [60], may be used. A key feature of iterative methods is that they only require matrix-vector products, rather than explicit matrices. In other words, the linear operator of Eq. (7) is never computed and stored, only its action on vectors is required. To further hasten calculations, the dimension of the basis N, and therefore the cutoff magnitude of G_m, is selected such that a discrete Fourier transform (DFT) can be used to transform between planewave and spatial representations [56]. This does of course restrict the lattice unit cells to those which can be represented by rectangular prisms, due to the uniform spatial discretization required for a DFT. Ultimately, the matrix-vector products are computed in $O(n \log n)$ time using the technique described in [56], the curl operations become cross products in wave vector space ($O(n)$) and the inverse permittivity is applied locally ($O(n)$) once a DFT ($O(n \log n)$) transforms solutions between wavevector and physical space. The limiting factor becomes the two DFTs required for every matrix-vector product, although this calculation is quite efficient using the Fastest Fourier Transform in the West (FFTW) library [61]. To further improve the computational efficiency, the calculations are parellelized; the curls computed in wavevector space are embarrassingly parallel, while the DFTs performed by FFTW

are parallelized along one spatial dimension. A further advantage of the planewave basis is the availability of an accurate preconditioner [56], which is crucial for rapid convergence of LOBPCG.

An open-source planewave expansion library, Maxwell Photonic Bands (MPB), is available, although the method was implemented from scratch in this work to facilitate the design sensitivity calculations. The only difference between the implementation used to generate the results in this paper and that described in [56] is the lack of permittivity smoothing here. This omission eases the sensitivity calculations, and any inaccuracies incurred by it are mitigated by refining the computational grid [56], which is done in our nested mesh refinement. Further, the generalization of permittivity smoothing to handle anisotropic materials [56] was not required here.

2.3. Photonic bandgap calculation

To predict the presence of a complete photonic bandgap, the propagating modes must be computed from Eq. (7) for all unique wave vectors κ in reciprocal lattice space [1]. Fortunately, the region of nonredundant wave vectors, i.e. the irreducible Brillouin zone (IBZ) denoted as $\mathcal{B} : \mathbb{R}^3$, is reduced for lattices possessing domain symmetries [62]. This work will consider, without loss of generality, orthorhombic crystals, due to their ease of computational implementation. $\mathcal{B}$ is simply a square, i.e. 1/4th of the unit cell cross-section, in 2D and an octant, i.e. 1/8th of the unit cell, in 3D with orthorhombic symmetry [62], cf. Figure 1. To be clear, all design parameterizations are 3D, but they will be optimized for both complete 2D bandgaps, where κ is restricted to lie in the $x - y$ plane, and complete 3D bandgaps where no such restriction is enforced. Note that a topology optimization scheme to overcome the difficulty of non-rectangular symmetry cells has been developed [34], although it is not needed here.
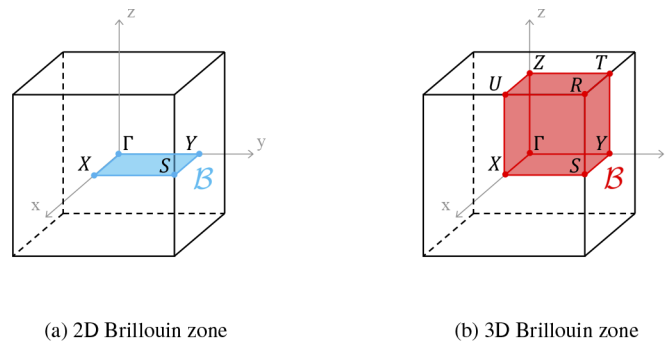

(a) 2D Brillouin zone   (b) 3D Brillouin zone

Fig. 1. Irreducible Brillouin zones of unit cells with orthorhombic symmetry

The bandgap-midgap ratio between bands m and $m + 1$ is defined as

$$\theta_\omega = \frac{\Delta\omega}{\bar{\omega}} = \frac{\min\,\omega_{m+1}(\kappa) - \max\,\omega_m(\kappa)}{\frac{1}{2}\left(\min\,\omega_{m+1}(\kappa) + \max\,\omega_m(\kappa)\right)} \quad \forall \kappa \in \mathcal{B}, \tag{12}$$

where $\mathcal{B}$ is the appropriate irreducible Brillouin zone for the unit cell. The bandgap-midgap ratio's scale invariance makes it preferable over the absolute bandgap $\Delta\omega$ as a performance metric [1]. Indeed, invariance ensures the lattice size of a photonic crystal may be scaled to shift the bandgap, i.e. $\bar{\omega}$, without changing θ_ω. Note that θ_ω lacks differentiability from both the min/max operations and the potential for degenerate eigenvalues. These matters are addressed in Sec. 2.4.3. Figure 2 displays an example dispersion plot with a gap between modes $m = 3$ and $m + 1 = 4$, with the gray region denoting the bandgap $[\omega_1, \omega_2]$. Notice that only the edges of $\mathcal{B}$, denoted as $\partial\mathcal{B}$, are included in the plot, since band extrema are rarely located on the interior of $\mathcal{B}$

[34]. A statistical analysis of band extrema location discovered that this assumption works best for crystals with high symmetry [63]. Therefore, only $\partial \mathcal{B}$ is included in our objective function calcluation, but a full $\mathcal{B}$ discretization is computed *a posteriori* to ensure our designs truly have a bandgap.
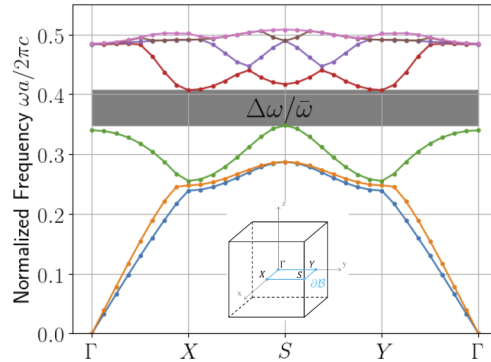

Fig. 2. Representative dispersion plot for 2D $\kappa \in \partial \mathcal{B}$

Previous work suggested that optimizing

$$\theta_\lambda = \frac{\min \lambda_{m+1}(\kappa) - \max \lambda_m(\kappa)}{\frac{1}{2}(\min \lambda_{m+1}(\kappa) + \max \lambda_m(\kappa))} \quad \forall \kappa \in \partial \mathcal{B}, \tag{13}$$

where $\lambda = \left(\frac{\omega}{c}\right)^2$, is equivalent to optimizing Eq. (12) and results in better optimization performance [34]. Thus, all optimizations in this work use Eq. (13) as the objective function, however, dispersion plots will be presented displaying the normalized frequency $\frac{\omega a}{2\pi c}$ for a unit cell of size a, as is customary. Fortunately, the eigenvalue computations at each wavevector are embarrassingly parallel, producing efficient strong scaling of the Brillouin scan. Note that the dependence of θ_λ on design variables was omitted here, but will be reconsidered in Sec. 2.4.4.

2.4. Topology optimization

Topology optimization distributes material in a design domain to generate optimal structures [17]. Traditionally, the structure Ω is defined by a material indicator function $\chi : \mathbb{R}^h \to \{0, 1\}$ such that $\Omega = \{x \in \mathbb{R}^h \mid \chi(x) = 1\}$, i.e. indicator values of $\chi(x) = 1$ correspond to the presence of solid material and values of $\chi(x) = 0$ correspond to void space at location x. The resulting integer programming problem is ill-posed due to a lack of inherent length scale [64]. A restriction method, such as a perimeter constraint [65], provides a minimum length scale and results in a well-posed problem, but the difficulty of the integer programming formulation remains. Convexifying the design space by replacing the binary-valued material indicator with the continuous volume fraction $\nu : \mathbb{R}^h \to [0, 1]$ enables the use of efficient, gradient-based optimization algorithms, and the use of alternative restriction techniques, such as a slope constraint [66] and filtration [67]. Unfortunately, designs now contain intermediate material regions where $\nu(x) \in (0, 1)$, as opposed to purely void and solid regions where $\nu(x) = 0$ and $\nu(x) = 1$, respectively. Often times additional measures must be taken to limit the extent of the intermediate volume fraction regions. For example, employing a material interpolation scheme [68–70] that penalizes the stiffness-to-weight ratio at locations where $\nu(x) \in (0, 1)$ in conjunction with a mass constraint [18] works well in the standard compliance topology optimization problems. Fortunately, photonic bandgap optimization problems are well-posed whence they do not require length scale control and when replacing the indicator function with the volume fraction to convexify the design space they naturally tend to binary designs, alleviating the need for material penalization schemes.

2.4.1. Design parameterization

The design variables d_i in this work are the values of the volume fraction over a set of "voxels" which discretize the unit cell, and thus the volume fraction field is piece-wise uniform. As mentioned in Sec. 2.3, orthorhombic symmetry will be considered thoughout this work. This symmetry is enforced by allowing only the design in the symmetry cell to change freely, i.e. by allowing each voxel in the symmetry cell to have its own volume fraction. The full unit cell design is obtained by mapping the symmetry cell design throughout the unit cell, as demonstrated by the 2D example in Fig. 3.
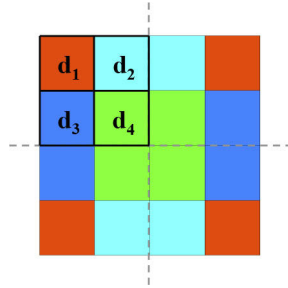

Fig. 3. Example 2D design parameterization with orthorhombic symmetry, dashed lines mark symmetry planes

The well-known "checker-boarding" phenomenon often observed in structural topology optimization due to the inherent ill-posedness of the problem [64] is naturally avoided when optimizing bandgap structures, as evidenced by numerical results [34,35,37,38] and theoretical considerations [40,41]. Nonetheless, a filtration technique is employed here to impose a minimum length scale to ensure manufacturability and to hasten convergence of the optimization. To this end, the "cone filter" presented in [67] is extended to periodic unit cells with imposed domain symmetry. The filter works by replacing the volume fraction with a smoothed, i.e. filtered, volume fraction field $\tilde{v}$. In this way a highly oscillatory v, which yields small scale features, is replaced by a smooth $\tilde{v}$. The cone filter uses a mesh-independent weighting of volume fraction that varies linearly with distance r such that

$$\tilde{v}(\boldsymbol{x}) = \int_{B_r(\boldsymbol{x})} K(\boldsymbol{x} - \boldsymbol{y}) \, v(\boldsymbol{y}) \, \mathrm{d}\Omega_y \tag{14}$$

where K is the linear kernel and $B_r(\boldsymbol{x})$ is a ball of radius r centered at $\boldsymbol{x}$. Upon discretization we obtain the discrete filtered volume fractions of the voxels $\tilde{d}_i$ from

$$\tilde{d}_i = W_{ij} d_j, \tag{15}$$

where

$$W_{ij} = \frac{\tilde{W}_{ij}}{\sum_i \tilde{W}_{ij}}, \tag{16}$$

with

$$\tilde{W}_{ij} = \max\left(0, \ r - ||\boldsymbol{x}_i - \boldsymbol{x}_j||_2\right), \tag{17}$$

in which $\boldsymbol{x}_i$ is the centroid of voxel i and $|| \cdot ||_2$ is Euclidean distance. Note that the entries are "normalized" such that the sum of each row of $\boldsymbol{W}$ equals 1. Extending this filter to a periodic domain is achieved by computing $||\boldsymbol{x}_i - \boldsymbol{x}_j||_2$ as the minimum distance between $\boldsymbol{x}_i$ and all valid positions of $\boldsymbol{x}_j$ in neighboring unit cells. Further, domain symmetry is accounted for by computing W_{ij} at each voxel centroid in the unit cell, and then adding all contributions to their

corresponding voxel in the symmetry cell. These operations are demonstrated in 2D by Fig. 4 and easily generalize to 3D. Note that the symmetry operations lead to "double counting" the contribution of elements near a symmetry cell boundary. The minimum length scale is roughly enforced by the selection of r; smaller values allow for smaller features.
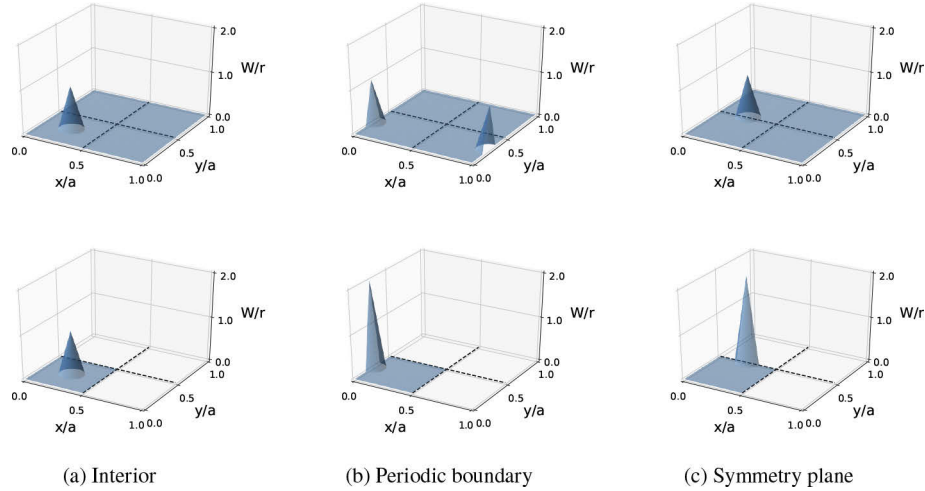

(a) Interior   (b) Periodic boundary   (c) Symmetry plane

Fig. 4. Cone filter kernel $B_r(x) K(x - y)$ in 2D at various locations displayed in unit cell (top row) and symmetry cell (bottom row), dashed lines mark symmetry planes

2.4.2. Material interpolation

The solid phase volume fraction v affects the design as we use a linear material interpolation scheme for the relative permittivity while solving Eq. (7), i.e.

$$\varepsilon_r(x) = \varepsilon_0 + \tilde{v}(x) (\varepsilon_1 - \varepsilon_0), \tag{18}$$

where $\tilde{v}(x)$ is the filtered volume fraction so that the relative permittivity $\varepsilon_r(x)$ is that of the solid phase ε_1 when $\tilde{v}(x) = 1$ and that of the vacuum ε_0 when $\tilde{v}(x) = 0$, respectively. Previous work has reported that when designing for maximal bandgaps using this interpolation, optimal structures naturally tend to binary designs, i.e. $\tilde{v}(x) \approx 0$ or $\tilde{v}(x) \approx 1$, in both 2D [10] and 3D [34,38] applications, which follows intuition that high dielectric contrast leads to large bandgaps [1]. Convergence to binary structures was also observed in 2D phononic bandgap topology optimization [35]. In agreement with previous works, our optimized structures naturally converged to binary designs eliminating the need for a nonlinear material interpolation scheme.

2.4.3. Objective function smoothing

As previously mentioned, the bandgap-midgap ratio defined in Eq. (13) lacks the differentiability required to use gradient-based optimization solvers, such as the method of moving asymptotes (MMA) [71] or interior point methods, e.g. IPOPT [72]. There are two distinct sources of non-smoothness, but fortunately both can be alleviated with the same technique. Repeated, or degenerate, eigenvalues are a well known issue plaguing eigenvalue optimization problems [26]. Consider a symmetric, matrix A which depends on design variables d. If A has eigenpairs (λ, ϕ) which satisfy $A\phi = \lambda\phi$, the simple eigenvalues λ can be differentiated with respect to a particular

design parameter d_i as [26]

$$\frac{\mathrm{D}\lambda}{\mathrm{D}d_i} = \frac{1}{\boldsymbol{\phi}^T\boldsymbol{\phi}} \boldsymbol{\phi}^T \frac{\partial \boldsymbol{A}}{\partial d_i}\boldsymbol{\phi}, \tag{19}$$

where D denotes a total derivative and ∂ denotes a partial derivative. It becomes clear that Eq. (19) admits infinitely many solutions in the case of repeated eigenvalues, wherein there are an infinite number of valid eigenvectors $\boldsymbol{\phi}$ in a hyper-plane with dimensionality equal to the eigenvalue's multiplicity [26]; as such, the derivative is not defined. Albeit, it is possible to compute directional eigenvalue derivatives using the generalized gradient method [30], however, most nonlinear programming algorithms cannot work with directional derivatives.

It was proven that a symmetric polynomial of eigenvalues, λ_i for $i = 1, \ldots, n$, is a differentiable function, provided the set is isolated, i.e. $\ldots \le \lambda_0 < \lambda_1 \le \ldots \le \lambda_n < \lambda_{n+1} \le \ldots$ [51]. Inspection of the inequalities reveals that the symmetric polynomial remains smooth even if the isolated set contains degenerate eigenvalues. This is an extremely powerful result that can be used to remove the eigenvalue degeneracy issue from topology optimization. Specifically, a "*p*-norm" function defined on a vector $\boldsymbol{a}$ as

$$l_p(\boldsymbol{a}) = \left(\sum_i a_i^p\right)^{\frac{1}{p}}, \tag{20}$$

is a symmetric polynomial raised to an exponent. We note that the usual *p*-norm absolute value is ignored since we assume $a_i > 0$. Thus, approximating the mode m eigenvalue

$$\lambda_m(\boldsymbol{\kappa}) \approx \tilde{\lambda}_m(\boldsymbol{\kappa}) = \left(\sum_{i=1}^{m} \lambda_i(\boldsymbol{\kappa})^p\right)^{\frac{1}{p}}, \tag{21}$$

and the mode $m + 1$ eigenvalue

$$\lambda_{m+1}(\boldsymbol{\kappa}) \approx \tilde{\lambda}_{m+1}(\boldsymbol{\kappa}) = \left(\sum_{i=m+1}^{M} \lambda_i(\boldsymbol{\kappa})^{-p}\right)^{\frac{1}{-p}}, \tag{22}$$

where M is sufficiently large to capture all repeated eigenvalues of $\lambda_{m+1}(\boldsymbol{\kappa})$, yields smooth approximations provided $\lambda_m \ne \lambda_{m+1}$. Although $\lambda_m \ne \lambda_{m+1}$ is not guaranteed across the entire design domain, the assumption is justified since it must hold for our bandgap designs. Note that our smooth approximation further requires $\lambda_M \ne \lambda_{M+1}$, although any errors introduced if $\lambda_M = \lambda_{M+1}$ are negligible in the sensitivity analysis provided $\lambda_M^p \gg \lambda_{m+1}^p$. So clearly, the accuracy of Eqs. (21) and (22) is also controlled by the value of p, which must be sufficiently large to obtain a reasonable approximation, although poor optimization convergence is likely if p is too large. Substituting the approximations from Eqs. (21) and (22) into Eq. (13) produces

$$\theta_\lambda \approx \frac{\min \tilde{\lambda}_{m+1}(\boldsymbol{\kappa}) - \max \tilde{\lambda}_m(\boldsymbol{\kappa})}{\frac{1}{2}\left(\min \tilde{\lambda}_{m+1}(\boldsymbol{\kappa}) + \max \tilde{\lambda}_m(\boldsymbol{\kappa})\right)} \quad \forall \boldsymbol{\kappa} \in \partial \mathcal{B}, \tag{23}$$

which effectively alleviates the non-differentiability caused by eigenvalue degeneracy.

Equation (23) still lacks the necessary differentiability for gradient-based optimization due to the min/max operations. Equation (20) is therefore applied to smoothly approximate these

operations as

$$\max \tilde{\lambda}_m(\boldsymbol{\kappa}) \approx \tilde{\lambda}_1 = \left(\sum_{i=1}^{N_{\boldsymbol{\kappa}}} \tilde{\lambda}_m(\boldsymbol{\kappa}_i)^p \right)^{\frac{1}{p}} \qquad \boldsymbol{\kappa} \in \partial \mathcal{B} \tag{24}$$

$$\min \tilde{\lambda}_{m+1}(\boldsymbol{\kappa}) \approx \tilde{\lambda}_2 = \left(\sum_{i=1}^{N_{\boldsymbol{\kappa}}} \tilde{\lambda}_{m+1}(\boldsymbol{\kappa}_i)^{-p} \right)^{\frac{1}{-p}} \qquad \boldsymbol{\kappa} \in \partial \mathcal{B}, \tag{25}$$

where $N_{\boldsymbol{\kappa}}$ is the number of wave vectors chosen along $\partial \mathcal{B}$. Substituting Eqs. (24) and (25) into Eq. (23) reveals the differentiable bandgap-midgap ratio approximation

$$\theta_\lambda \approx \tilde{\theta}_\lambda = \frac{\tilde{\lambda}_2 - \tilde{\lambda}_1}{\frac{1}{2} \left(\tilde{\lambda}_2 + \tilde{\lambda}_1 \right)}. \tag{26}$$

2.4.4. Sensitivity analysis

As previously noted, the bandgap-midgap ratio will be influenced by the solid material volume fraction, which is parameterized by a set of design variables $\boldsymbol{d}$. Thus, the derivative of Eq. (26) is required with respect to each design parameter d_i to perform efficient gradient-based optimization. The chain rule is used to compute each derivative according to

$$\frac{\mathrm{D}\tilde{\theta}_\lambda}{\mathrm{D}d_i} = \frac{\partial \tilde{\theta}_\lambda}{\partial \tilde{\lambda}_1} \frac{\mathrm{D}\tilde{\lambda}_1}{\mathrm{D}d_i} + \frac{\partial \tilde{\theta}_\lambda}{\partial \tilde{\lambda}_2} \frac{\mathrm{D}\tilde{\lambda}_2}{\mathrm{D}d_i}, \tag{27}$$

which requires

$$\frac{\partial \tilde{\theta}_\lambda}{\partial \tilde{\lambda}_1} = \frac{-4\tilde{\lambda}_2}{\left(\tilde{\lambda}_2 + \tilde{\lambda}_1 \right)^2} \tag{28}$$

and

$$\frac{\partial \tilde{\theta}_\lambda}{\partial \tilde{\lambda}_2} = \frac{4\tilde{\lambda}_1}{\left(\tilde{\lambda}_2 + \tilde{\lambda}_1 \right)^2}. \tag{29}$$

To compute the total derivatives $\frac{\mathrm{D}\tilde{\lambda}_1}{\mathrm{D}d_i}$ and $\frac{\mathrm{D}\tilde{\lambda}_2}{\mathrm{D}d_i}$, the derivative of Eq. (20) for a design dependent input vector $\boldsymbol{a}$ is evaluated via the chain rule as

$$\frac{\mathrm{D}l_p(\boldsymbol{a})}{\mathrm{D}d_i} = \sum_j \frac{\partial l_p(\boldsymbol{a})}{\partial a_j} \frac{\mathrm{D}a_j}{\mathrm{D}d_i}, \tag{30}$$

where

$$\frac{\partial l_p(\boldsymbol{a})}{\partial a_j} = \left(\sum_k a_k^p \right)^{\frac{1}{p}-1} a_j^{p-1}. \tag{31}$$

Eq. (30) is applied to evaluate the derivatives of Eqs. (24) and (25), and then again applied to evaluate the derivatives of Eqs. (21) and (22), which are required for the chain rule applications. As discussed in Sec. 2.4.3, we see that the errors in the derivative computations of Eqs. (22) and (25) due to $\lambda_M = \lambda_{M+1}$ are small if $\lambda_M^p \gg \lambda_{m+1}^p$. Note that in the smooth minimum approximations we employ a negative exponent, i.e. $-p$, and therefore a_j^{-p-1}, cf. Equation (31), becomes very small for large a_j and large p.

At this point, only the derivatives of the eigenvalues $\frac{\mathrm{D}\lambda}{\mathrm{D}d_i}$ remain. Despite the possibility for degenerate eigenvalues, Eq. (19) may be utilized with assured smoothness through the use of symmetric polynomials [51]. Computing Eq. (19) requires the discretized matrix-vector product $\frac{\partial A}{\partial d_i} \boldsymbol{\phi}$, which may be computed using the matrix-free technique described in Sec. 2.2 with the

caveat of multiplying by $\frac{\mathrm{D}}{\mathrm{D}d_i}\left(\frac{1}{\varepsilon_r(x)}\right)$, rather than $\frac{1}{\varepsilon_r(x)}$, after transforming the solution via DFT to physical space. This technique may be feasible for a small number of design variables, but will likely become computationally prohibitive for the large number of design variables used in topology optimization. Fortunately, a more efficient technique to compute the $\frac{\mathrm{D}\lambda}{\mathrm{D}d_i}$ exists when design variables have local support, which is the case in topology optimization. To see this we examine a weak formulation of the eigenvalue problem, which is obtained by integrating the dot product of Eq. (5) with an arbitrary weighting function w over the computational domain Ω as

$$\int_\Omega w(x) \cdot \left(\nabla \times \left(\frac{1}{\varepsilon_r(x)}\nabla \times H(x)\right)\right)\ \mathrm{d}\Omega = \int_\Omega w(x) \cdot \lambda H(x)\ \mathrm{d}\Omega. \tag{32}$$

The vector identity

$$\nabla \cdot (a \times b) = b \cdot (\nabla \times a) - a \cdot (\nabla \times b) \tag{33}$$

is then applied to move a derivative from the solution field H to the weighting function w. Indeed, substituting $a = \frac{1}{\varepsilon_r}\nabla \times H$ and $b = w$ converts Eq. (32) to

$$\int_\Omega \nabla \cdot \left(\left(\frac{1}{\varepsilon_r(x)}\nabla \times H(x)\right) \times w(x)\right)\ \mathrm{d}\Omega + \int_\Omega \left(\frac{1}{\varepsilon_r(x)}\nabla \times H(x)\right) \cdot (\nabla \times w(x))\ \mathrm{d}\Omega$$
$$= \lambda \int_\Omega w(x) \cdot H(x)\ \mathrm{d}\Omega, \tag{34}$$

and applying the divergence theorem to the first integral yields

$$\int_{\partial\Omega} \hat{n}(x) \cdot \left(\left(\frac{1}{\varepsilon_r(x)}\nabla \times H(x)\right) \times w(x)\right)\ \partial\Omega + \int_\Omega \left(\frac{1}{\varepsilon_r(x)}\nabla \times H(x)\right) \cdot (\nabla \times w(x))\ \mathrm{d}\Omega$$
$$= \lambda \int_\Omega w(x) \cdot H(x)\ \mathrm{d}\Omega, \tag{35}$$

where $\hat{n}$ is the outward facing normal vector. Finally, the periodic boundary condition annihilates the boundary integral leaving the eigenvalue problem of finding the admissible H such that

$$\int_\Omega \frac{1}{\varepsilon_r(x)}(\nabla \times H(x)) \cdot (\nabla \times w(x))\ \mathrm{d}\Omega = \lambda \int_\Omega w(x) \cdot H(x)\ \mathrm{d}\Omega. \tag{36}$$

for all admissible w. Here we see that upon discretization, Eq. (36) is of the form $AH = \lambda H$. Taking the variation of Eq. (36) yields

$$\int_\Omega \delta\left(\frac{1}{\varepsilon_r(x)}\right)(\nabla \times H(x)) \cdot (\nabla \times w(x))\ \mathrm{d}\Omega + \int_\Omega \frac{1}{\varepsilon_r(x)}(\nabla \times \delta H(x)) \cdot (\nabla \times w(x))\ \mathrm{d}\Omega$$
$$= \delta\lambda \int_\Omega w(x) \cdot H(x)\ \mathrm{d}\Omega + \lambda \int_\Omega w(x) \cdot \delta H(x)\ \mathrm{d}\Omega. \tag{37}$$

where we use Eq. (36) to cancel the δw terms, since δw is admissible. Assigning the arbitrary weighting function $w = H$ yields

$$\int_\Omega \delta\left(\frac{1}{\varepsilon_r(x)}\right)(\nabla \times H(x)) \cdot (\nabla \times H(x))\ \mathrm{d}\Omega + \int_\Omega \frac{1}{\varepsilon_r(x)}(\nabla \times \delta H(x)) \cdot (\nabla \times H(x))\ \mathrm{d}\Omega$$
$$= \delta\lambda \int_\Omega H(x) \cdot H(x)\ \mathrm{d}\Omega + \lambda \int_\Omega H(x) \cdot \delta H(x)\ \mathrm{d}\Omega, \tag{38}$$

which reduces to

$$\int_\Omega \delta\left(\frac{1}{\varepsilon_r(x)}\right)(\nabla \times H(x)) \cdot (\nabla \times H(x))\ \mathrm{d}\Omega = \delta\lambda \int_\Omega H(x) \cdot H(x)\ \mathrm{d}\Omega, \tag{39}$$

where we again used the fact that Eq. (36) holds for the admissible $w = \delta H$. Solving for $\delta\lambda$ leaves the expression

$$\delta\lambda = \left(\int_{\Omega} H(x) \cdot H(x) \ d\Omega \right)^{-1} \int_{\Omega} (\nabla \times H(x)) \cdot \delta\left(\frac{1}{\varepsilon_r(x)} \right) (\nabla \times H(x)) \ d\Omega. \tag{40}$$

Upon evaluating $\delta\lambda$ with respect to a variation in the voxel i filtered volume fraction $\tilde{d}_i$, we obtain

$$\frac{D\lambda}{D\tilde{d}_i} = \left(\int_{\Omega} H(x) \cdot H(x) \ d\Omega \right)^{-1} \int_{\Omega} (\nabla \times H(x)) \cdot \frac{D}{D\tilde{d}_i}\left(\frac{1}{\varepsilon_r(x)} \right) (\nabla \times H(x)) \ d\Omega. \tag{41}$$

We notice here that upon discretization Eq. (41) is of the same form as Eq. (19). Equation (41) is advantageous for the eigenvalue derivative computation because the second integral is only non-zero over the voxel i corresponding to filtered design variable $\tilde{d}_i$. The key to using this technique with a plane wave basis is to compute $\nabla \times H$ in wave vector space and then obtain the physical vectors from a DFT. Also, many eigenvalue solvers scale the eigenvectors such that $\int_{\Omega} H(x) \cdot H(x) \ d\Omega = 1$, so this global integral may not need to be computed. The execution time will likely be many orders of magnitude (problem size dependent) faster than treating local design variables as if they were global and using Eq. (19). Finally, the chain rule is completed by evaluating the derivative of the permittivity in the voxel i, i.e.

$$\frac{D}{D\tilde{d}_i}\left(\frac{1}{\varepsilon_r} \right) = \frac{-1}{\varepsilon_r^2} \frac{D\varepsilon_r}{D\tilde{d}_i}, \tag{42}$$

where $\frac{D\varepsilon_r}{D\tilde{d}_i}$ is obtained by differentiating Eq. (18) as

$$\frac{D\varepsilon_r}{D\tilde{d}_i} = (\varepsilon_1 - \varepsilon_0). \tag{43}$$

The above procedure is used to obtain derivatives of $\tilde{\theta}_\lambda$ with respect to filtered volume fractions $\tilde{d}_i$, however, we must supply derivatives with respect to the design variables d_j to the NLP algorithm. Differentiating Eq. (15) provides the required relation

$$\frac{D\tilde{d}_i}{Dd_j} = W_{ij}, \tag{44}$$

so that we may finally compute

$$\frac{D\tilde{\theta}_\lambda}{Dd_j} = \frac{D\tilde{\theta}_\lambda}{D\tilde{d}_i} \frac{D\tilde{d}_i}{Dd_j} = \frac{D\tilde{\theta}_\lambda}{D\tilde{d}_i} W_{ij}. \tag{45}$$

The steps outlined above are used to compute the sensitivity of the objective function with respect to the volume fraction of each voxel in the unit cell. However, since only the volume fractions of voxels in the symmetry cell are free design variables, the sensitivity contributions from each voxel in the unit cell must be appropriately allocated to its corresponding voxel in the symmetry cell. For a 3D design parameterization with orthorhombic symmetry, the sensitivity contributions from the 8 symmetry reflected voxels in the unit cell are summed to compute the sensitivity with respect to the volume fraction of their corresponding symmetry cell voxel.

2.4.5. Nested mesh refinement

The full dispersion analysis required to predict a photonic bandgap becomes a computationally expensive endeavor as the solution basis and the discretization of $\partial\mathcal{B}$ are both refined. Parallel

calculations speed the analysis considerably; each eigenvalue calculation can be parallelized along one-dimension of the unit cell, while the scan of the Brillouin zone boundary $\partial \mathcal{B}$ is embarrassingly parallel, as noted in Sec. 2.3. Unfortunately, the calculation wall clock time can still exceed a few hours on an advanced supercomputer when the number of design variables exceeds 10^3, despite leveraging multi-level parallelism. To further reduce the computational cost of our 3D studies, a nested mesh refinement technique is developed.

Nonlinear progamming is used to solve the bound-constrained optimization problem

$$
\begin{aligned}
\underset{\vec{d}}{\text{maximize}} \quad & \tilde{\theta}_\lambda \\
\text{subject to} \quad & 0 \le d_j \le 1,
\end{aligned}
\tag{46}
$$

where $\tilde{\theta}_\lambda$ is defined in Eq. (26). Note that modes m, and thus $m + 1$, are selected *a priori*. We use a reduced space approach whereby the Maxwell equations are strictly enforced and accounted for in the sensitivity analysis presented in Sec. 2.4.4. The discrete representation of the symmetry cell begins as a coarse grid of only 2^3 voxels whose volume fractions are randomly assigned. This symmetry cell is subsequently reflected within the unit cell to enforce orthorhombic symmetry, corresponding to an analysis grid of 4^3 voxels. We then solve the topology optimization problem on this coarse grid. Upon convergence, the optimized design field is uniformly refined, cf. Figure 5, and a subsequent optimization begins from the converged coarse design. Note that this uniform refinement does not change the physical structure (sans filtering), but rather subdivides each voxel in the symmetry cell into 8 sub-voxels, whose volume fractions serve as design variables in the refined topology optimization problem. This process continues until the design grid has been sufficiently refined to resolve the desired feature size. It is also recommended to refine the grid at least twice to achieve reasonable bandgap predictions.
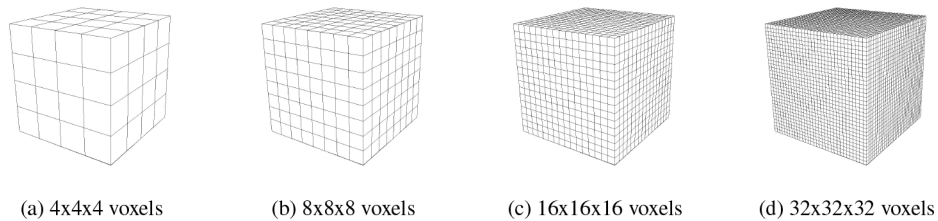

(a) 4x4x4 voxels　　　(b) 8x8x8 voxels　　　(c) 16x16x16 voxels　　　(d) 32x32x32 voxels

Fig. 5. Uniformly refined unit cells

2.4.6. Post-processing

A post-processing analysis of the optimized design is performed for two reasons. The first is to interpret structures consisting purely of binary designs, and the second is to verify that the photonic bandgap predicted by considering only a discrete set of vectors κ on the Brillouin zone boundary $\partial \mathcal{B}$ is valid, i.e. that there are no wave vectors κ in the interior of $\mathcal{B}$ that will propagate with a frequency that lies in the bandgap. Therefore, designs are thresholded by rounding volume fractions to 0 or 1 such that voxels with intermediate volume fractions are removed and a binary structure remains. The entire irreducible Brillouin zone, i.e. the interior $\mathcal{B}$ *and* boundary $\partial \mathcal{B}$, is then considered by analyzing this binary structure at wave vectors κ from a uniformly spaced sampling to ensure that the predicted bandgaps are valid. This uniform sampling of $\mathcal{B}$ consisted of 11 samples per spatial dimensions, for a total of 1331 κ vectors. Fortunately, in all cases it was determined that the bandgap predicted by scanning the boundary $\partial \mathcal{B}$ was consistent with the bandgap predicted by scanning the entirety of $\mathcal{B}$. Further, the thresholded binary designs were often superior due to their sharper material contrast.

3. Results and discussion

3.1. Complete 2D bandgaps

The techniques described in Sec. 2. were applied to design 3D photonic crystals with complete 2D bandgaps. The relevant numerical parameters that remain constant thoughout this work are listed in Tab. 1. Structures with complete 2D bandgaps were designed by solving Eq. (46) with a 2D Brillouin zone such that $\partial \mathcal{B}$ is restricted to the $x - y$ plane, with a κ spacing of $\frac{2\pi}{8a}$. The first 8 possible gaps are considered, which begin between modes 2 and 3, since it is impossible to split modes 1 and 2 due their shared frequency of 0 at $\kappa = \mathbf{0}$. Note that M from Eq. (22) was selected such that 4 modes above the desired bandgap were computed, i.e. $M = m + 4$; this was sufficient in all cases to ensure differentiability of the p-norm, i.e. such that $\lambda_M^p \gg \lambda_{m+1}^p$. The filter radius r is selected *a priori* to limit feature size. Since r remains constant throughout the mesh refinements, the filter will have no effect until the voxel spacing Δx is smaller than r. Thus, the filter is "activated" at some point in the mesh refinement process. It was observed in practice that proper selection of r is paramount to achieving a bandgap structure. If r is too large then the design may not have the necessary freedom to create a bandgap, while if r is too small fine-scale features appear which are difficult to fabricate and optimizations take longer to converge. Specifically, the best results were obtained when r was selected in the range $\frac{a}{32} < r < \frac{a}{8}$, which always activates the filter before attempting the fourth optimization on the 32^3 voxel unit cell.

Table 1. Numerical Parameters

Description	Symbol	Value
void space relative permittivity	ε_0	1.0
solid material relative permittivity	ε_1	13.0
p-norm exponent	p	30
$\partial \mathcal{B}$ discretization	$\Delta\kappa$	$\frac{2\pi}{8a}$
eigenvalue tolerance	ϵ_λ	1e-5
NLP tolerance	ϵ_θ	1e-5

The initial and optimal volume fractions, along with their corresponding dispersion plots, for each step in the nested mesh refinement process are displayed in Fig. 6 for a structure exhibiting a 2D bandgap between modes 5 and 6. A filter radius of $r = 0.05a$ was selected. Note that the initial design for the coarsest resolution is random. It is apparent that each design in the nested mesh refinement process converges to roughly the same design as its predecessor. Geometric features that cannot be fully resolved by the coarse discretizations appear with intermediate volume fractions, but as the mesh is refined the designs tend toward binary, aside from the interface smearing due to the filtering. As seen here, once a filter radius is selected to enforce a minimum feature size, mesh refinements must occur until features with those sizes can be fully resolved, resulting in a binary design without strictly enforcing this binary behavior.

An interesting observation is that the design performance is often over-estimated on the coarse grid. Fortunately, this effect is reduced in each mesh refinement step as the simulations become more accurate. Not surprisingly, design performance is decreased when the filter is activated, cf. Figure 6. Note that this filter activation refers to a uniform mesh refinement that causes $\Delta x < r$, i.e. the design variables do not change. This follows intuition that a photonic bandgap prefers stark material contrast, which is reduced when the filter smears the material interface. Despite these effects, the increased design freedom due to mesh refinement usually improves performance.

The post-processsed unit cell and its corresponding periodic photonic crystal are illustrated in Fig. 7, along with the dispersion plot from this binary design. Images of post-processed, i.e. binary, designs throughout this work display only the solid material to better examine the
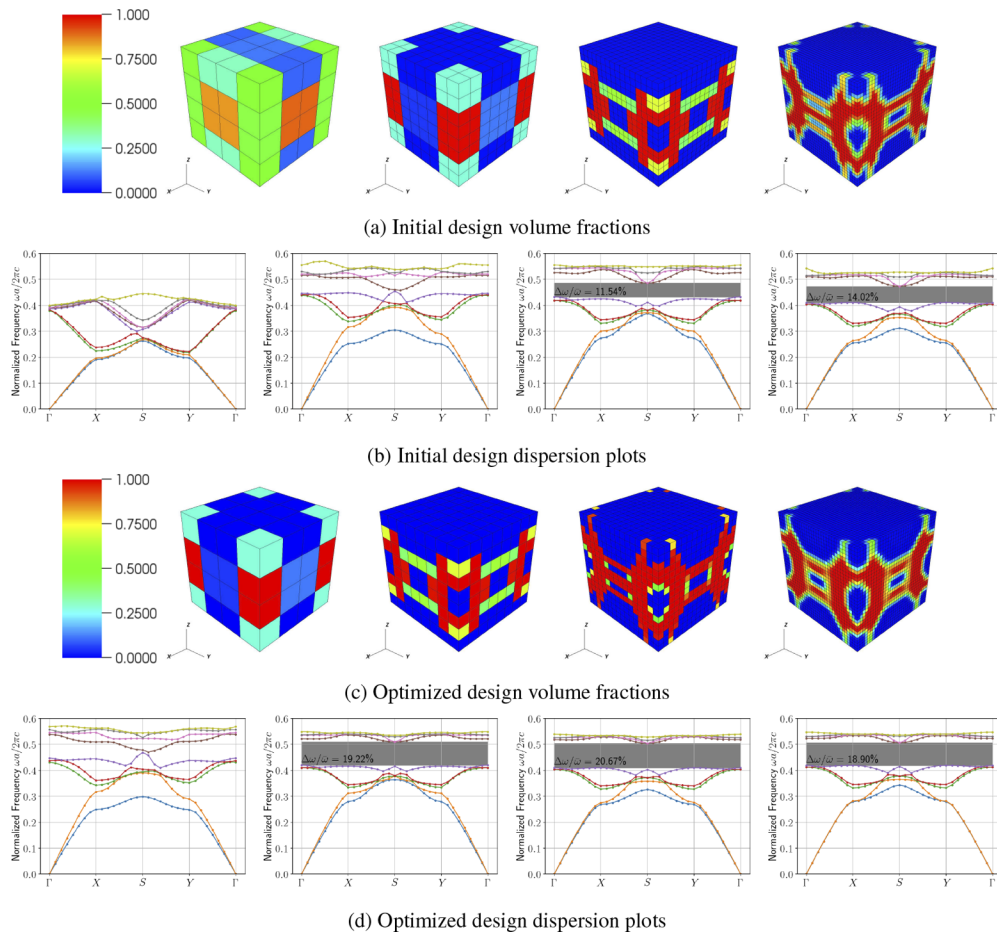
(a) Initial design volume fractions

(b) Initial design dispersion plots

(c) Optimized design volume fractions

(d) Optimized design dispersion plots

Fig. 6. Nested mesh refinement for complete 2D bandgap

structure. The full $\mathcal{B}$ scan described in Sec. 2.4.6 was also performed to verify the bandgap performance predicted by the $\partial\mathcal{B}$ scan.
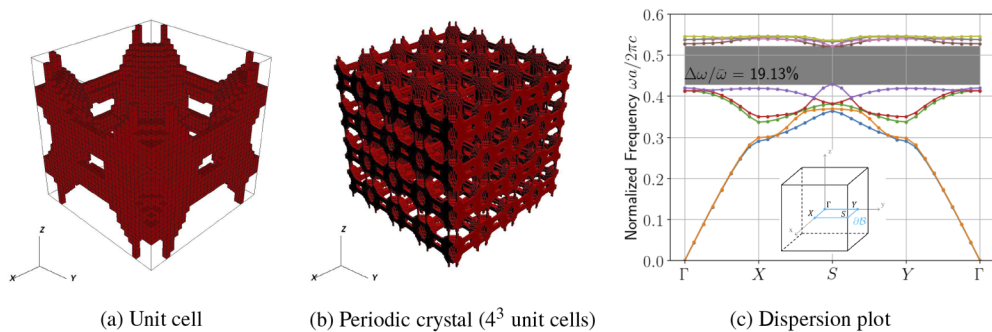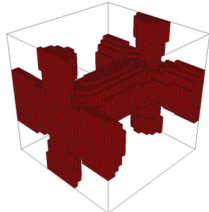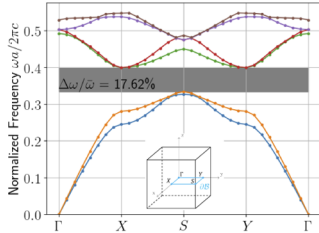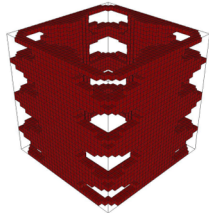

(a) Unit cell

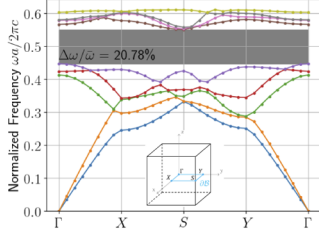(b) Periodic crystal (4^3 unit cells)

(c) Dispersion plot
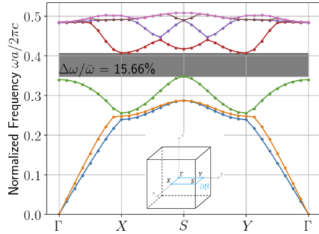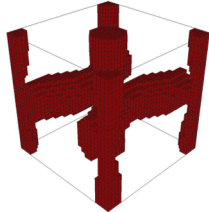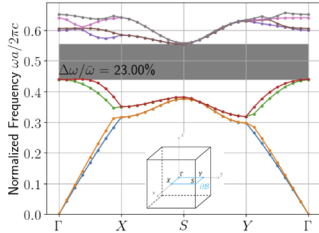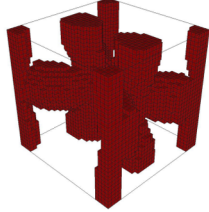
Fig. 7. Post-processed photonic crystal with complete 2D bandgap between modes 5-6

Tables. 2 and 3 present the best 2D photonic bandgap structures obtained from 30 random initial designs for each of the first 8 possible bandgaps. The post-processed performance of

the best structures are also presented, along with the percentage of random initial designs that yield bandgap structures. Although a formal study was not performed, our experience shows that increasing the κ spacing above $\frac{2\pi}{8a}$ results in a smaller percentage of initial designs yielding bangdap structures.
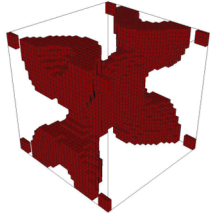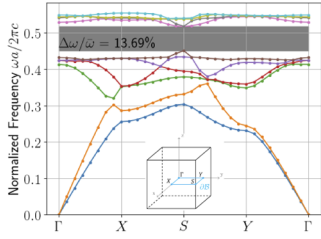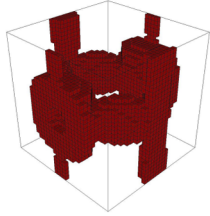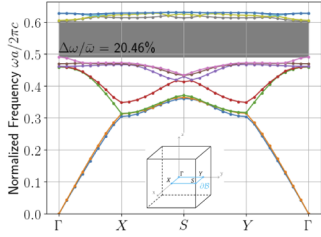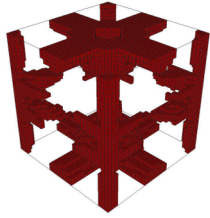
Table 2. Summary of complete 2D bandgap optimizations (gaps 1-4)

Modes (m-m+1)	Best structure	Dispersion plot	Bandgap probability
2-3			$\frac{26}{30} = 86.7\%$
3-4			$\frac{10}{30} = 33.3\%$
4-5			$\frac{7}{30} = 23.3\%$
5-6			$\frac{27}{30} = 90.0\%$

3.2. Complete 3D Bandgaps

Designing photonic crystals with complete 3D bandgaps is considerably more difficult than with complete 2D bandgaps. Similar to the 2D bandgap study, 30 random initial designs were used for each of the first 8 possible bandgaps, yet only 3 unique photonic crystals with complete 3D bandgaps were obtained using the same $r = 0.05a$ filter radius. The optimal designs at each mesh resolution for the three successful 3D bandgap designs are displayed in Figs. 8–10, along
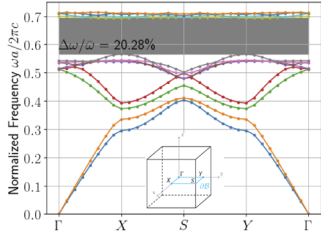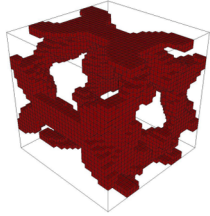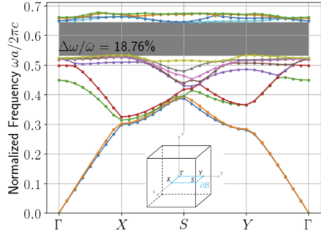
Table 3. Summary of complete 2D bandgap optimizations (gaps 5-8)

Modes (m-m+1)	Best structure	Dispersion plot	Bandgap probability
6-7			$\frac{1}{30} = 3.3\%$
7-8			$\frac{3}{30} = 10.0\%$
8-9			$\frac{12}{30} = 40.0\%$
9-10			$\frac{2}{30} = 6.67\%$

with their post-processed unit cells and dispersion plots. Many of the random initial designs converged to the same optimal structure; specifcally, 19/30 converged to the design pictured in Fig. 8 for $m = 2$, 6/30 converged to the design pictured in Fig. 9 for $m = 5$, and 7/30 converged to the design pictured in Fig. 10 for $m = 6$.
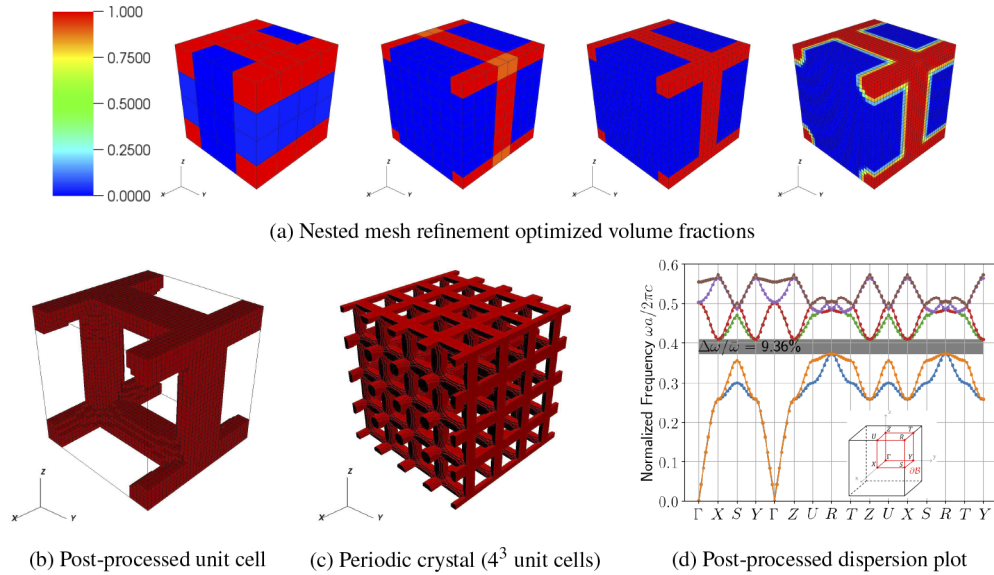

(a) Nested mesh refinement optimized volume fractions

(b) Post-processed unit cell   (c) Periodic crystal (4^3 unit cells)   (d) Post-processed dispersion plot

Fig. 8. Photonic crystal with complete 3D bandgap between modes 2-3

(a) Nested mesh refinement optimized volume fractions

(b) Post-processed unit cell   (c) Periodic crystal (4^3 unit cells)   (d) Post-processed dispersion plot
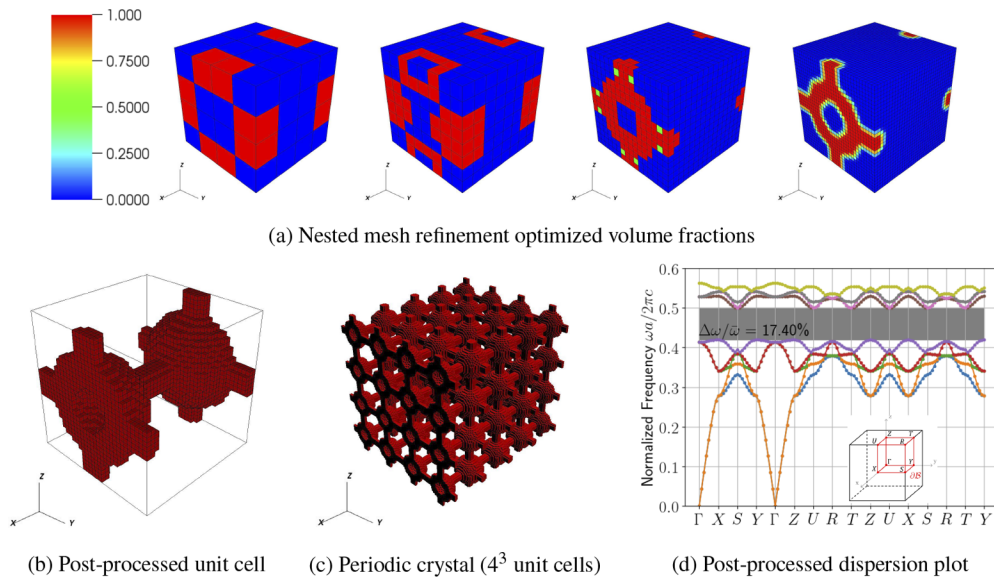
Fig. 9. Photonic crystal with complete 3D bandgap between modes 5-6

It is interesting to note that the design in Fig. 9 is very similar to a previously published optimal design [34]. This further validates the presented implementation. Also, it lends confidence to the possibility that global optimum are obtainable by the nested mesh refinement strategy,
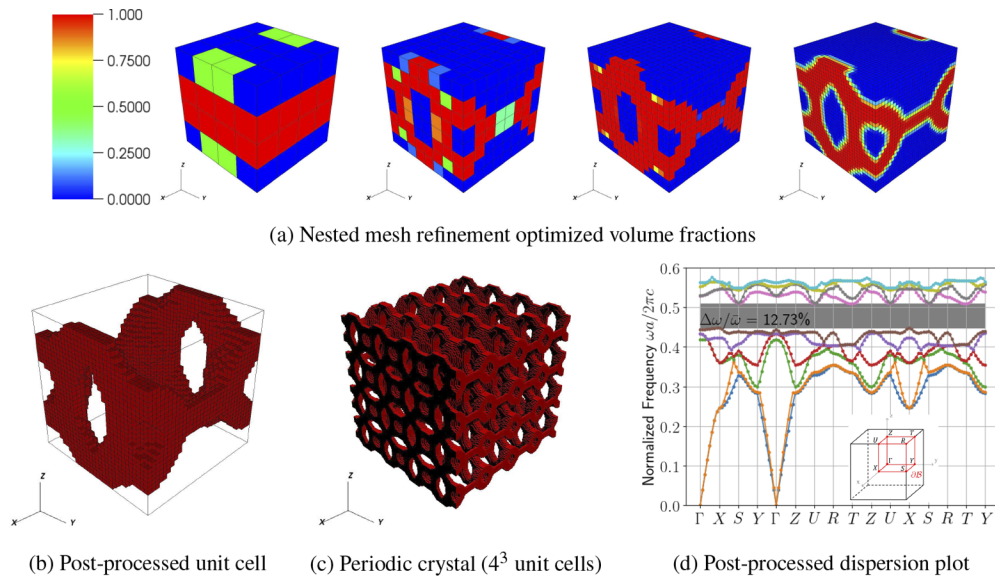
(a) Nested mesh refinement optimized volume fractions

(b) Post-processed unit cell

(c) Periodic crystal (4^3 unit cells)

(d) Post-processed dispersion plot

Fig. 10. Photonic crystal with complete 3D bandgap between modes 6-7

since [34] used a fixed grid throughout the optimization. Additionally, it is promising that the semidefinite programming reformulation in [34] produced the same result as the interior point NLP algorithm employed here, implying a consistency between methods and further justifying our lack of permittivity smoothing. An advantage of the proposed method over the semidefinite programming method is the ability to easily incorporate other design parameterizations and nonlinear constraints.

The structure displayed in Fig. 10 has not, to the knowledge of the authors, been previously published. Although the performance does not exceed that of known structures, it is important to have the capability to find new bandgap structures that may be more robust to defects, or have better multi-functional properties. Extending the presented framework to other unit cell symmetries or attempting to split higher modes with a more refined mesh could very well produce novel bandgap structures with better performance than what has previously been produced.

3.3. Computational expense

Photonic crystal design for complete bandgap is a computationally expensive endeavor. The discrete eigenvalue analyses resulting from the fully-vectorial formulation consist of real and imaginary parts with two DOF per voxel, producing a discrete system with $4n^3$ DOF for an analysis grid with n voxels along each spatial dimension of the unit cell. For context, the eigenvalue problems derived from the 32^3 grids have $131,072$ DOF. Further, it is generally required to compute the lowest ≈ 10 eigenvalues, and it was necessary to analyze 32 and 92 κ vectors along $\partial\mathcal{B}$ to design for complete 2D and 3D bandgaps, respectively. The NLP algorithm generally converged within ≈ 300 design iterations on each of the four nested mesh refinement grids for a total of $\approx 38,400$ and $\approx 110,400$ eigenvalue analyses when designing for complete 2D and 3D bandgaps, respectively. The 2D bandgap optimizations were solved on 512 processors, while the 3D bandgap optimizations were solved on 1,472 processors, both requiring ≈ 4 hours of execution time. Nested mesh refinement significantly reduced the number of iterations required, although the time savings were not computed since convergence to a bandgap structure from a random initial design at the 32^3 grid resolution could not be obtained within the 24 hour time limit of the Lawrence Livermore National Laboratory (LLNL) computing cluster used throughout.

4.   Conclusions

Topology optimization has successfully been applied to design 3D photonic crystals with complete 2D and 3D bandgaps. A technique for smoothing the bandgap-midgap ratio utilizes the symmetric architecture of a "*p*-norm" function to remove nondifferentiability issues caused by eigenvalue degeneracies and min/max operations, allowing the usual gradient-based nonlinear programming solvers to be employed.

Nested mesh refinement alleviates the computational burden from the high dimensional design spaces required to optimize 3D bandgap structures. Optimal designs naturally converge to binary structures after sufficient mesh refinement, and a periodic cone filter is successfully applied to enforce a minimum length scale in the optimized designs.

The presented framework is general enough to design photonic crystals with any domain symmetry, provided the unit cell is a rectangular prism. Three-dimensional photonic crystals with orthorhombic symmetry were designed with complete 2D bandgaps between each of the first 8 possible mode pairs, and complete 3D bandgaps between 3 different mode pairs. The sizes of the bandgaps presented here are competitive with the largest described in the literature.

Disclosures. The authors declare no conflicts of interest.

Data Availability. No data were generated or analyzed in the presented research.

References

1. J. Joannopoulous, S. Johnson, J. Winn, and R. Meade, *Photonic Crystals: Molding the Flow of Light* (Princeton University, 2008), 2nd ed.
2. Lord Rayleigh Sec. R. S., "Xvii. on the maintenance of vibrations by forces of double frequency, and on the propagation of waves through a medium endowed with a periodic structure," The London, Edinburgh, Dublin Philos. Mag. J. Sci. **24**(147), 145–159 (1887).
3. E. Yablonovitch, "Inhibited spontaneous emission in solid-state physics and electronics," Phys. Rev. Lett. **58**(20), 2059–2062 (1987).
4. S. John, "Strong localization of photons in certain disordered dielectric superlattices," Phys. Rev. Lett. **58**(23), 2486–2489 (1987).
5. E. Yablonovitch and T. Gmitter, "Photonic band structure: the face-centered-cubic case," Phys. Rev. Lett. **63**(18), 1950–1953 (1989).
6. E. Yablonovitch, "Photonic band-gap structures," J. Opt. Soc. Am. B **10**(2), 283–295 (1993).
7. E. Özbay, A. Abeyta, G. Tuttle, M. Tringides, R. Biswas, C. T. Chan, C. M. Soukoulis, and K. M. Ho, "Measurement of a three-dimensional photonic band gap in a crystal structure made of dielectric rods," Phys. Rev. B **50**(3), 1945–1948 (1994).
8. K. Ho, C. Chan, and C. Soukoulis, "Existence of a photonic gap in periodic dielectric structures," Phys. Rev. Lett. **65**(25), 3152–3155 (1990).
9. J. Smajic, C. Hafner, and D. Erni, "Design and optimization of an achromatic photonic crystal bend," Opt. Express **11**(12), 1378–1384 (2003).
10. J. Jensen and O. Sigmund, "Systematic design of photonic crystal structures using topology optimization: low-loss waveguide bends," Appl. Phys. Lett. **84**(12), 2022–2024 (2004).
11. S. A. Rinne, F. García-Santamaría, and P. V. Braun, "Embedded cavities and waveguides in three-dimensional silicon photonic crystals," Nat. Photonics **2**(1), 52–56 (2008).
12. T. Ma, A. B. Khanikaev, S. H. Mousavi, and G. Shvets, "Guiding electromagnetic waves around sharp corners: Topologically protected photonic transport in metawaveguides," Phys. Rev. Lett. **114**(12), 127401 (2015).
13. J. Joannopoulous, "The almost magical world of photonic crystals," in *1999 IEEE LEOS Annual Meeting Conference Proceedings. LEOS'99. 12th Annual Meeting. IEEE Lasers and Electro-Optics Society 1999 Annual Meeting (Cat. No.99CH37009)*, vol. 1 (1999), pp. 232–233.
14. J. Geremia, J. Williams, and H. Mabuchi, "nverse-problem approach to designing photonic crystals for cavity qed experiments," Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top. **66**(6), 066606 (2002).
15. M. B. Dühring, O. Sigmund, and T. Feurer, "Design of photonic bandgap fibers by topology optimization," J. Opt. Soc. Am. B **27**(1), 51–58 (2010).
16. G. Allaire, *Shape Optimization by the Homogenization Method* (Springer-Verlag New York, Inc., 2002).

17. M. Bendsøe and N. Kikuchi, "Generating optimal topologies in structural design using a homogenization method," Comput. Methods Appl. Mech. Eng. **71**(2), 197–224 (1988).

18. M. P. Bendsøe and O. Sigmund, *Topology Optimization: Theory, Methods, and Applications* (Springer, 2004), 2nd ed.

19. O. Sigmund, "Materials with prescribed constitutive parameters: an inverse homogenization problem," Int. J. Solids Struct. **31**(17), 2313–2329 (1994).

20. M. Osanov and J. Guest, "Topology optimization for architected materials design," Annu. Rev. Mater. Res. **46**(1), 211–233 (2016).

21. Y. Ding, Z. Liu, C. Qiu, and J. Shi, "Matameterial with simultaneously negative bulk modulus and mass density," Phys. Rev. Lett. **99**(9), 093904 (2007).

22. K. Wojciechowski, "Two-dimensional isotropic system with a negative poisson ratio," Phys. Lett. A **137**(1-2), 60–64 (1989).

23. D. Smith, W. Padilla, D. Vier, S. Nemat-Nasser, and S. Schultz, "Composite medium with simultaneously negative permeability and permittivity," Phys. Rev. Lett. **84**(18), 4184–4187 (2000).

24. D. A. Pawlak, "Metamaterials and photonic crystals – potential applications for self-organized eutectic micro- and nanostructures," Scientia Plena **4**, 014801 (2008).

25. M. Burger, S. Osher, and E. Yablonovitch, "Inverse problem techniques for the design of photonic crystals," IEICE Transactions on Electronics **E87C**, 258–265 (2004).

26. A. Seyranian, E. Lund, and N. Olhoff, "Multiple eigenvalues in structural optimization problems," Struct. Optim. **8**(4), 207–227 (1994).

27. M. Doosje, B. J. Hoenders, and J. Knoester, "Photonic bandgap optimization in inverted fcc photonic crystals," J. Opt. Soc. Am. B **17**(4), 600–606 (2000).

28. S. J. Cox and D. C. Dobson, "Maximizing band gaps in two-dimensional photonic crystals," SIAM J. Appl. Math. **59**(6), 2108–2120 (1999).

29. S. J. Cox and D. C. Dobson, "Band structure optimization of two-dimensional photonic crystals in h-polarization," J. Comput. Phys. **158**(2), 214–224 (2000).

30. S. Cox, "The generalized gradient at a multiple eigenvalue," J. Funct. Analysis **133**(1), 30–40 (1995).

31. C. Kao, S. Osher, and E. Yablonovitch, "Maximizing band gaps in two-dimensional photonic crystals using level set methods," Appl. Phys. B **81**(2-3), 235–244 (2005).

32. T. Felici and H. W. Engl, "On shape optimization of optical waveguides using inverse problem techniques," Inverse Probl. **17**(4), 1141–1162 (2001).

33. H. Men, N. Nguyen, R. Freund, P. Parrilo, and J. Peraire, "Bandgap optimization of two-dimensional photonic crystals using semidefinite programming and subspace methods," J. Comput. Phys. **229**(10), 3706–3725 (2010).

34. H. Men, K. Lee, R. Freund, J. Peraire, and S. Johnson, "Robust topology optimization of three-dimensional photonic-crystal band-gap structures," Opt. Express **22**(19), 22632–22648 (2014).

35. O. Sigmund and J. Jensen, "Systematic design of phononic band–gap materials and structures by topology optimization," Philos. Trans. R. Soc., A **361**(1806), 1001–1019 (2003).

36. H. Dong, Y. Wang, T. Ma, and X. Su, "Topology optimization of simultaneous photonic and phononic bandgaps and highly effective phoxonic cavity," J. Opt. Soc. Am. B **31**(12), 2946–2955 (2014).

37. L. Shen, Z. Ye, and S. He, "Design of two-dimensional photonic crystals with large absolute band gaps using a genetic algorithm," Phys. Rev. B **68**(3), 035109 (2003).

38. F. Meng, B. Jia, and X. Huang, "Topology-optimized 3d photonic structures with maximal omnidirectional bandgaps," Adv. Theory Simul. **1**(12), 1870032 (2018).

39. O. Sigmund, "On the usefulness of non-gradient approaches in topology optimization," Struct. Multidiscip. Optim. **43**(5), 589–596 (2011).

40. O. Sigmund and K. Hougaard, "Geometric properties of optimal photonic crystals," Phys. Rev. Lett. **100**(15), 153904 (2008).

41. S. Li, H. Lin, F. Meng, D. Moss, X. Huang, and V. Jia, "On-demand design of tunable complete photonic band gaps based on bloch mode analysis," Sci. Rep. **8**(1), 14283 (2018).

42. J. Jensen and O. Sigmund, "Topology optimization for nano-photonics," Laser Photonics Rev. **5**(2), 308–321 (2011).

43. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," Nat. Photonics **12**(11), 659–670 (2018).

44. I. L. Lyubchanskii, N. N. Dadoenkova, M. I. Lyubchanskii, E. A. Shapovalov, and T. Rasing, "Magnetic photonic crystals," J. Phys. D: Appl. Phys. **36**(18), R277–R287 (2003).

45. V. I. Belotelov and A. K. Zvezdin, "Magneto-optical properties of photonic crystals," J. Opt. Soc. Am. B **22**(1), 286–292 (2005).

46. M. Inoue, R. Fujikawa, A. Baryshev, A. Khanikaev, P. B. Lim, H. Uchida, O. Aktsipetrov, A. Fedyanin, T. Murzina, and A. Granovsky, "Magnetophotonic crystals," J. Phys. D: Appl. Phys. **39**(8), R151–R161 (2006).

47. M. Nur-E-Alam, M. Vasiliev, and K. Alameh, "ano-structured magnetic photonic crystals for magneto-optic polarization controllers at the communication-band wavelengths," Opt. Quantum Electron. **41**(9), 661–669 (2009).

48. V. Kotov, V. Shavrov, M. Vasiliev, K. Alameh, M. Nur-E-Alam, and D. Balabanov, "Properties of magnetic photonic crystals in the visible spectral region and their performance limitations," Photonics Nanostructures - Fundamentals Appl. **28**, 12–19 (2018).

49. K. C. Huang, E. Lidorikis, X. Jiang, J. D. Joannopoulos, K. A. Nelson, P. Bienstman, and S. Fan, "Nature of lossy bloch states in polaritonic photonic crystals," Phys. Rev. B: Condens. Matter Mater. Phys. **69**(19), 195111 (2004).

50. A. Raman and S. Fan, "Photonic band structure of dispersive metamaterials formulated as a hermitian eigenvalue problem," Phys. Rev. Lett. **104**(8), 087401 (2010).

51. J. Gravesen, A. Evgrafov, and D. M. Nguyen, "On the sensitivities of multiple eigenvalues," Struct. Multidiscip. Optim. **44**(4), 583–587 (2011).

52. F. H. Clarke, *Optimization and Nonsmooth Analysis* (Society for Industrial and Applied Mathematics, 1990).

53. J. Jin, *Theory and Computation of Electromagnetic Fields* (John Wiley & Sons, Inc., 2015), 2nd ed.

54. F. Bloch, "Über die quantenmechanik der elektronen in kristallgittern," Eur. Phys. J. A **52**(7-8), 555–600 (1929).

55. P. Kuchment, *Floquet theory for partial differential equations* (Springer, 1993).

56. S. G. Johnson and J. D. Joannopoulos, "Block-iterative frequency-domain methods for maxwell's equations in a planewave basis," Opt. Express **8**(3), 173–190 (2001).

57. J. Jin, *The Finite Element Method in Electromagnetics* (John Wiley & Sons, Inc., 2014), 3rd ed.

58. J. Nédélec, "Mixed finite elements in $\mathbb{R}^3$," Numer. Math. **35**(3), 315–341 (1980).

59. "MFEM: Modular finite element methods library," mfem.org.

60. A. Knyazev, "Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method," SIAM J. Sci. Comput. **23**(2), 517–541 (2001).

61. M. Frigo and S. G. Johnson, "The design and implementation of FFTW3," Proc. IEEE **93**, 216–231 (2005).

62. W. Setyawan and S. Curtarolo, "High-throughput electronic band structure calculations: challenges and tools," Comput. Mater. Sci. **49**(2), 299–312 (2010).

63. F. Maurin, C. Claeys, E. Deckers, and W. Desmet, "Probability that a band-gap extremum is located on the irreducible brillouin-zone contour for the 17 different plane crystallographic lattices," Int. J. Solids Struct. **135**, 26–36 (2018).

64. O. Sigmund and J. Petersson, "Numerical instabilities in topology optimization: A survey on procedures dealing with checkerboards, mesh-dependencies and local minima," Struct. optimization **16**(1), 68–75 (1998).

65. R. B. Haber, C. S. Jog, and M. P. Bendsøe, "A new approach to variable-topology shape design using a constraint on perimeter," Struct. optimization **11**(1-2), 1–12 (1996).

66. J. Peterssson and O. Sigmund, "Slope constrained topology optimization," Numer. Methods Eng. **41**(8), 1417–1434 (1998).

67. T. E. Bruns and D. Tortorelli, "Topology optimization of non-linear elastic structures and compliant mechanisms," Comput. Methods Appl. Mech. Eng. **190**(26-27), 3443–3459 (2001).

68. M. P. Bendsøe, "Optimal shape design as a material distribution problem," Struct. optimization **1**(4), 193–202 (1989).

69. M. Stolpe and K. Svanberg, "An alternative interpolation scheme for minimum compliance topology optimization," Struct. Multidiscip. Optim. **22**(2), 116–124 (2001).

70. M. P. Bendsøe and O. Sigmund, "Material interpolation schemes in topology optimization," Arch. Appl. Mech. **69**(9-10), 635–654 (1999).

71. K. Svanberg, "The method of moving asymptotes—a new method for Struct. optimization," Int. J. for Numer. Methods Eng. **24**(2), 359–373 (1987).

72. A. Wächter and L. Biegler, "On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming," Math. Program. **106**(1), 25–57 (2006).