SIAM J. APPL. MATH.
Vol. 73, No. 6, pp. 2188–2202

A NUMERICAL APPROACH FOR DEFECT MODES LOCALIZATION IN AN INHOMOGENEOUS MEDIUM*

YIQI GU† AND XIAOLIANG CHENG†

Abstract. Some optical design problems arise from the study of photonic bandgap structure, including defect modes localization, that is, computing the optimal dielectric property to highly localize particular eigenfunctions of a Dirichlet model problem. The steepest descent method has been studied for this problem. In this paper, we present a new approach for the defect modes localization. Rather than focusing on the original objective and optimizing the structure along the gradient, a variant of the original problem is put forward with its corresponding method. Although the original problem and the variant presented in this work are not equivalent, our method is shown to solve both of them in numerical experiments. Furthermore, the algorithm in this paper can restart the optimization if the original gradient descent method gets stuck during the iteration.

Key words. mode localization, defect modes, spectrum problem, optimal design, photonic bandgap

AMS subject classifications. 65K10, 82D25, 49M07

DOI. 10.1137/120883566

1. Introduction.

In recent years, there has been some research involving photonic crystals (first introduced in [8, 15]) and the bandgap phenomenon, in which the wave whose frequency is in some band cannot propagate in this medium. The phenomenon is related to loss mechanisms in optical and mechanical systems, thus leading to a series of optimal problems about the medium. For example, we can derive the eigenvalue problem from wave equations, then set up the objective whose variables are the eigenpairs and search for the optimal solution in an admissible set. Cox and Dobson [13, 14] have considered the optimization of bandgap in two-dimensional periodic structures composed of two given dielectric materials, including the E-polarization and H-polarization cases. Another level set method can be used in this optimization [5]; Lipton, Shipman, and Venakides [12] have optimized the electromagnetic resonant properties in periodic photonic crystal slabs, from the relation between resonance and transmission; Kao and Santosa [6] have derived the maximization of quality factor from wave equations in an inhomogeneous medium, whose index of refraction is the design variable. This optimization is related to a nonlinear eigenvalue problem, and the quality factor is defined as the ratio between the real part of the complex eigenfrequency and the imaginary part. Numerical results demonstrate that the optimal solution is a piecewise constant function, which is analytically proved by Karabash [9]; Heider et al. [10] have optimized scattering resonances in micro- and nano-scale components, thus decreasing the radiative loss, which is the magnitude of the imaginary part of scattering resonances; Osting [3] considered the optimization of two spectral functions from the Dirichlet–Laplacian eigenvalue problem. The variable domains are

*Received by the editors July 5, 2012; accepted for publication (in revised form) August 22, 2013; published electronically December 5, 2013.
http://www.siam.org/journals/siap/73-6/88356.html
†Department of Mathematics, Zhejiang University, 866 Yuhangtang Road, Hangzhou, Zhejiang Province, 310058, People's Republic of China (yiqig@uw.edu, xiaoliangcheng@zju.edu.cn). The second author's research was supported by Key Project of the Major Research Plan of NSFC (91130004) and NSFC (J1210038).

represented by Fourier-cosine coefficients and a BFGS quasi-Newton method is used to search for the optimal region.

This work focuses on the problem of defect modes localization, previously discussed by Dobson and Santosa [7]. In the two-dimensional Dirichlet eigenvalue problem, we are to find material properties to localize the defect modes, which are certain eigenpairs of the spectral equation. A simple model problem with a Dirichlet boundary can be utilized to replace the original complicated problems, and the goal is to lead the eigenfunction highly localized. In [7], a gradient descent method is proposed to deal with this optimization. Denoting the objective as J and the considered design as b, the method evaluates the gradient $g = \frac{\delta J}{\delta b}$ by an adjoint approach and updates b along the direction. Some other optimization problems of photonic crystals are also solved by the gradient method or generalized gradient method [5, 6, 10, 13, 14].

The optimal design problem is considered from a new perspective in this work. We search for the designs aiming at getting desired eigenvectors, instead of getting lower objective $J(b)$. A variant of the original problem is presented, as well as the corresponding formula and algorithms. Different from the idea in [7], a preknown objective (the desired eigenvector) is set and approximated. Although the new optimization presented by us is not equivalent to the original one, it is found in numerical experiments that the original objective $J(b)$ is indeed decreased, as well as the objective of the new problem (see section 5).

When testing the original gradient algorithm presented in [7] in some specific examples, we notice that the process may get stuck. That is, the objective $J(b)$ stops decreasing and even increases at some stage instead of converging to a well-optimized solution. One reason for this phenomenon may be the process falling into a local minimum extrema. Another probable reason is the restrictive operator worked on the variable b making its components bounded in the admissible set after each update, which "pulls back" the variable to a poor position. However, our modified Algorithm 2 (see section 4) can deal with this case much better and seldom stagnates at an early stage. A corresponding example is shown in section 5.

The outline of this paper is as follows. In section 2 we introduce the background of the model problem, including the original formulation and a new variant. A restricted form of the model problem is specifically presented for analysis, and the variant is deduced from the Rayleigh quotient property of symmetric matrices. We consider the problem from the relationship between the variation of the design and the eigenvector in section 3 and propose the numerical methods in section 4. In section 5 we show some examples to demonstrate the effects and features of the algorithm. A conclusion is given in section 6. The symbols in this paper are as in [7], and $\|\cdot\|$ stands for $\|\cdot\|_2$.

2. Problem description. As we know, the propagation of waves will be impeded in the medium with periodic structures which have certain bandgaps. In this case, a point defect can be introduced, leading to a spatially localized standing wave. This work aims at the optimization of the localized modes. We use $\epsilon_p(x)$ to characterize the dielectric property of the periodic medium. Having a photonic bandgap means the spectral problem

$$\Delta u + \omega^2 \epsilon_p(x) u = 0, \quad x \in \mathbb{R}^2,$$

has a gap in its continuous spectrum. We introduce a defect into the medium, that is, give a perturbation $\eta(x)$ with compact support to $\epsilon_p(x)$. It brings the form

$$\Delta u + \omega^2 (\epsilon_p(x) + \eta(x)) u = 0, \quad x \in \mathbb{R}^2. \tag{2.1}$$

If an eigenpair $(\omega, u(x))$ of (2.1) satisfies that ω falls into the bandgap of the periodic medium and $u(x)$ decays exponentially away from the defect, it is defined as a defect mode. More details about defect modes are given in [1, 2, 7, 11]. Different perturbations cause different defect modes, and we are to find $\eta(x)$ which brings highly localized eigenfunctions.

2.1. Model problem. The problem we consider is a simplified version, which is restricted in a domain with Dirichlet boundary. The model problem not only can overcome the difficulty of handling the original unbounded domain but also can consider $\epsilon(x)$ and $\eta(x)$ as a single variable. Moreover, in this simple version, we need not consider satisfying explicitly the conditions to create defect modes in a photonic bandgap structure. The simplified model shows the major property of the original problem, and the approaches solving it can be applied to the original one.

The Dirichlet eigenvalue problem is described as follows:

$$-\Delta u = \lambda \epsilon u \quad \text{in } \Omega, \tag{2.2a}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{2.2b}$$

where $u \in H_0^1(\Omega)$ is the modes; $\epsilon(x)$ is the dielectric coefficient of the medium, satisfying $0 < \epsilon_0 \leqslant \epsilon(x) \leqslant \epsilon_1 < \infty$; and Ω is a simply connected bounded domain in $\mathbb{R}^2$ with Lipschitz continuous boundary. We normalize u to keep unit energy in the domain, that is,

$$\int_\Omega \epsilon u^2 = 1. \tag{2.3}$$

The goal is to find a proper dielectric coefficient $\epsilon(x)$, to make the eigenfunctions u most localized. Therefore, we introduce the objective

$$J(\epsilon, u) = \int_\Omega w \epsilon u^2$$

to measure the degree of localization, where w is some given weight function satisfying $w(0,0) = 0$.

To make the problem well-posed, a global form and a local form are presented in [7]. However, what we consider is another restricted form of (2.2). Discretizing the operator $-\Delta$ to a corresponding matrix noted as A, $u(x)$ to a n-dimensional vector noted as u, and $\epsilon(x)$ to a corresponding diagonal matrix noted as $S(\epsilon)$, which multiplies the entries of u pointwise by the elements in $\epsilon(x)$, we get

$$Au = \lambda S(\epsilon) u. \tag{2.4}$$

It is assumed A and $S(\epsilon)$ are both symmetric and positive definite, and thus (2.4) can be rewritten as

$$AS^{-\frac{1}{2}}(S^{\frac{1}{2}}u) = \lambda S^{\frac{1}{2}}(S^{\frac{1}{2}}u).$$

Let $S^{-\frac{1}{2}} = B$, $S^{\frac{1}{2}}u = v$; then

$$BABv = \lambda v, \tag{2.5a}$$

$$\langle v, v \rangle = 1, \tag{2.5b}$$

where (2.5b) is from the discretization of (2.3) and $\langle \cdot, \cdot \rangle$ means the inner product of vectors. We construct a vector b using the diagonal elements of B and its entry $b^{(j)} = 1/\sqrt{\epsilon(x_j)}$, where x_j is the jth discretized point. We define an admissible set $\mathcal{A} = \{b \in \mathbb{R}^n : b_0 \leqslant b^{(j)} \leqslant b_1\}$, where $b_0 = 1/\sqrt{\epsilon_1}$ and $b_1 = 1/\sqrt{\epsilon_0}$, which are the bounds of b's entries; then the set $\mathcal{A}$ will be searched for the optimal designs.

Defining the objective under discretization as

$$J(b) = \frac{1}{2}\langle v(b), Wv(b)\rangle, \tag{2.6}$$

where, W is a prescribe symmetric weight matrix, $v(b)$ is some eigenvector of BAB with unit modulus, the unrestricted optimal design problem will be

$$\min_{b\in\mathcal{A}} J(b).$$

To be compatible with some restricted forms, such as the global problem and local problem described in [7], we discuss the following restricted form: set $b_0 \in \mathcal{A}$, $B_0 = \mathrm{diag}(b_0)$ and v_0 satisfies

$$B_0AB_0v_0 = \lambda_i v_0, \quad \langle v_0, v_0\rangle = 1,$$

where λ_i is the ith (single) eigenvalue of B_0AB_0. So v_0 is well-defined regardless of the sign; then the objective $J_i(b_0) = \frac{1}{2}\langle v_0, Wv_0\rangle$ is well-evaluated. Viewing B_0 as variable B, we let it vary continuously; then v_0 will vary weak-continuously as well, noted as v. (For the possible eigenvalues crossing, v may not always belong to the ith eigenvalue of BAB.) If v is always single, $J_i(b)$ will be a continuous functional with variable b. We need to evaluate

$$\min_{b\in\mathcal{A}} J_i(b). \tag{2.7}$$

If λ_0 is multiple, the initial vector v_0 should be appointed specifically from the multidimensional eigenspace.

It is easy to see $J_i(b)$ is related not only to b but also to b_0. How b_0 varies will affect $J_i(b)$, that is, b_0 can be changed into b through different ways, and thus the ith eigenpair of B_0AB_0 may be changed into different results, causing $J_i(b)$ to be not well-defined. For this reason, we should consider this problem in a practical view, i.e., find a way to modify b_0, thus minimizing $J_i = \frac{1}{2}\langle v_0, Wv_0\rangle$ as far as possible.

2.2. Variant of problem. We introduce the Rayleigh quotient of matrix W

$$R(W, x) = \frac{\langle x, Wx\rangle}{\langle x, x\rangle}, \quad x \in \mathbb{R}^n,$$

whose maximum and minimum can be evaluated from the Courant–Fisher min-max theorem. For simplicity, it is stated as follows.

THEOREM 2.1. *Suppose A is a real symmetric matrix; then the Rayleigh quotient $R(A, x)$, $x \in \mathbb{R}^n$, satisfies*

$$\lambda_1 = R(A, q_1) \leqslant R(A, x) \leqslant R(A, q_n) = \lambda_n,$$

where λ_1 and λ_n are the smallest and largest eigenvalue of A, and q_1 and q_n are the associated eigenvectors.

It is known from Theorem 2.1 $R(W,x) \geqslant R(W,v^*) = \lambda_{\min}(W)$, where v^* is the eigenvector with unit modulus belonging to the smallest eigenvalue of W. Then we have

$$\langle v^*, Wv^* \rangle = \min_{v \in \mathbb{R}^n} \langle v, Wv \rangle.$$

PROPOSITION 2.2. *J is defined as* (2.6), *v, $v^* \in \mathbb{R}^n$, $\|v\| = \|v^*\| = 1$; then*

$$|J(v) - J(v^*)| \leqslant \|W\| \cdot \min\{\|v + v^*\|, \|v - v^*\|\}.$$

Proof.

$$|J(v) - J(v^*)| = \left| \frac{1}{2}\langle v, Wv \rangle - \frac{1}{2}\langle v^*, Wv^* \rangle \right| = \frac{1}{2} \left| \langle v - v^*, W(v + v^*) \rangle \right|,$$

using the Cauchy–Schwarz inequality,

$$\begin{aligned} |J(v) - J(v^*)| &\leqslant \frac{1}{2}\|v - v^*\| \cdot \|W(v + v^*)\| \\ &\leqslant \frac{1}{2}\|v + v^*\| \cdot \|W\| \cdot \|v - v^*\| \leqslant \|W\| \cdot \|v - v^*\|. \end{aligned}$$

Similarly,

$$|J(v) - J(v^*)| \leqslant \|W\| \cdot \|v + v^*\|.$$

Thus the proposition is proven. □

From Proposition 2.2, it is natural to take v close to $\pm v^*$, making $|J(v) - J(v^*)|$ have a lower bound. The optimization (2.7) can be transformed to

$$\min_{b \in \mathcal{A}} \|v(b) - v^*\|. \tag{2.8}$$

Although (2.7) and (2.8) are not equivalent, we observe in numerical experiments that by approximating $v(b)$ to v^*, the function $J_i(b)$ is decreased simultaneously. In the next section, a numerical method is presented to solve problem (2.8). We first derive the relationship between the perturbation of b and v from (2.5), then describe the approach to make v approximate to v^*.

3. Method description. Rewrite (2.5) as

$$BABv = \langle v, BABv \rangle v,$$

where we suppose v is the unit eigenvector associated with the ith eigenvalue. Give b a small perturbation Δb, which causes a linearized response Δv in v; then we get

$$(vv^{\mathrm{T}} - I)(BA\mathrm{diag}(v) + \mathrm{diag}(ABv))\Delta b = (BAB - \lambda I - 2\lambda vv^{\mathrm{T}})\Delta v, \tag{3.1}$$

where $O(\Delta v^2)$ is ignored. The preceding equation shows if we want v to have a perturbation Δv, we can just add the perturbation Δb to b, which can be computed from (3.1). However, (3.1) with respect to unknown Δb has a factor $vv^{\mathrm{T}} - I$ in the coefficient, which is singular with rank $n - 1$. So Δb cannot be computed via matrix inversion. We choose to compute Δb through the normal equation of (3.1), thus obtaining a least square solution restricted to a $(n-1)$-dimensional subspace.

As the symmetry of BAB, we can do its eigenvalue decomposition $BAB = VDV^{\mathrm{T}}$, where V is orthogonal and D is diagonal formed by the eigenvalues of BAB. By multiplying each column of V by $vv^{\mathrm{T}} - I$, we derive

$$vv^{\mathrm{T}} - I = V\mathrm{diag}(-1, \ldots, 0^{(i)}, \ldots, -1)V^{\mathrm{T}}. \tag{3.2}$$

Denote $M \triangleq BA\mathrm{diag}(v) + \mathrm{diag}(ABv)$; then (3.1) can be rewritten as

$$\mathrm{diag}(-1, \ldots, 0^{(i)}, \ldots, -1)V^{\mathrm{T}}M\Delta b = V^{\mathrm{T}}(BAB - \lambda I - \lambda vv^{\mathrm{T}})\Delta v. \tag{3.3}$$

To derive the normal equation, we premultiply $\mathrm{diag}(-1, \ldots, 0^{(i)}, \ldots, -1)$ on both sides of (3.3), then obtain

$$\begin{aligned} &\mathrm{diag}(1, \ldots, 0^{(i)}, \ldots, 1)V^{\mathrm{T}}M\Delta b \\ &\quad = \mathrm{diag}(-1, \ldots, 0^{(i)}, \ldots, -1)V^{\mathrm{T}}(BAB - \lambda I - 2\lambda vv^{\mathrm{T}})\Delta v \\ &\quad = \mathrm{diag}(-1, \ldots, 0^{(i)}, \ldots, -1)V^{\mathrm{T}}(BAB - \lambda I)\Delta v. \end{aligned} \tag{3.4}$$

Notice the common factor $\mathrm{diag}(1, \ldots, 0^{(i)}, \ldots, 1)V^{\mathrm{T}}$ of (3.4) can be cancelled, leading a simplified equation

$$M\Delta b = -(BAB - \lambda I)\Delta v. \tag{3.5}$$

Therefore if Δb is a solution of (3.5), it also solves (3.4). We can just compute Δb from (3.5).

As the solution from (3.5) is restricted to a subspace of the original equation (3.1), we cannot expect that the computed perturbation Δb added to b will accurately trigger the given variation Δv of v. However, the distance between the true perturbation of v after adding Δb to b and the previously given Δv can be estimated as follows.

PROPOSITION 3.1. *Under the definitions above and with the assumption λ is a single nonzero eigenvalue, let $u, \Delta b^* \in \mathbb{R}^n$ satisfiy $M\Delta b^* = -(BAB - \lambda I)u$; then*

$$\|x(\Delta b^*) - u\|_2 = \min_{\Delta b \in \mathbb{R}^n} \|x(\Delta b) - u\|_2,$$

where $x(\Delta b)$ is the unique solution of equations

$$(BAB - \lambda I - 2\lambda vv^T)x = (vv^T - I)M\Delta b$$

with unknown x.

Proof. Let $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and λ is the ith eigenvalue, namely, λ_i. From the fact

$$BAB - \lambda I - 2\lambda vv^{\mathrm{T}} = V\mathrm{diag}(\lambda_1 - \lambda, \ldots, -2\lambda^{(i)}, \ldots, \lambda_n - \lambda)V^{\mathrm{T}}$$

and the singleness of λ we know $BAB - \lambda I - 2\lambda vv^{\mathrm{T}}$ is nonsingular. On the other hand, using (3.2), we have

$$M\Delta b^* = -(BAB - \lambda I)u,$$

so

$$\mathrm{diag}(1 \cdots 0^{(i)} \cdots 1)V^{\mathrm{T}}M\Delta b^* = \mathrm{diag}(\lambda_1 - \lambda, \ldots, 0^{(i)}, \ldots, \lambda_n - \lambda)V^{\mathrm{T}}(-u),$$

which is equivalent to

$$M^{\mathrm{T}}V\mathrm{diag}((\lambda_1-\lambda)^{-2},\ldots,0^{(i)},\ldots,(\lambda_n-\lambda)^{-2})V^{\mathrm{T}}M\Delta b^* = M^{\mathrm{T}}V\mathrm{diag}((\lambda_1-\lambda)^{-1},\ldots,0^{(i)},\ldots,(\lambda_n-\lambda)^{-1})V^{\mathrm{T}}(-u)$$

and so to

$$\widetilde{M}^{\mathrm{T}}\widetilde{M}\Delta b^* = \widetilde{M}^{\mathrm{T}}u,$$

where $\widetilde{M} = (BAB - \lambda I - 2\lambda vv^{\mathrm{T}})^{-1}(vv^{\mathrm{T}} - I)M$. Therefore, Δb^* is the least square solution of the equation $\widetilde{M}y = u$ with unknown y, that is,

$$\|\widetilde{M}\Delta b^* - u\|_2 = \min_{y\in\mathbb{R}^n}\|\widetilde{M}y - u\|_2. \tag{3.6}$$

After putting $x(\Delta b^*) = \widetilde{M}\Delta b^*$ into (3.6), we finish the proof. □

Now we examine the singularity of M, noticing the fact

$$\begin{aligned} M &= \mathrm{diag}(ABv) + BA\mathrm{diag}(v) \\ &= \mathrm{diag}(B^{-1}\lambda v) + BABB^{-1}\mathrm{diag}(v) \\ &= (\lambda I + BAB)B^{-1}\mathrm{diag}(v). \end{aligned} \tag{3.7}$$

It is straightforward that the condition number of $\lambda I + BAB$ is decided by the ratio between $\lambda + \lambda_{\max}(BAB)$ and $\lambda + \lambda_{\min}(BAB)$, and the condition number of B^{-1} is controlled by the element bounds b_0 and b_1. So in usual computations $(\lambda I + BAB)B^{-1}$ is well-conditioned, and the singularity of M is mainly decided by the ratio between the largest and smallest entries of v. If there exists one element of v very close to zero, M will be ill-conditioned.

4. Basic algorithm. From (3.5), it is easy to propose an algorithm to make the eigenvector belonging to the ith eigenvalue of BAB approach the objective v^*. Supposing the design in kth iteration, denoted as b_k, is just computed, we can next compute the eigenvector v_k (of B_kAB_k) and compute Δv_k according to some rules. Under the assumption that Δb_k can be determined exactly from the relation (3.1), we let $b_{k+1} \triangleq b_k + \Delta b_k$ and compute v_{k+1} as an eigenvector of $B_{k+1}AB_{k+1}$ which continuously varies from v_k; then it is a new approximation to v^* and satisfies $v_{k+1} = v_k + \Delta v_k$. Defining Δv_k is a crucial task in the algorithm, which requires the series $\{v_k\}$ to converge to v^* gracefully. For convenience, we can define $\Delta v_k \triangleq \omega(v^* - v_k)$, where ω is a small real factor to control the modulus of Δv_k so that formulations derived from perturbation analysis will be effective. However, considering the near orthogonality between v^* and v_k in most situations, we think this definition is too "abrupt" and the continuity of the shape of v_k is broken. In fact, we observe in our experiments that even though the objective $J_i(b)$ falls down a lot, v^* and v_k are still nearly orthogonal. Consequently, we search for some appropriate manners by which the series $\{v_k\}$ approaches v^* more "smoothly." Noticing the gradient of $J(v) = \frac{1}{2}\langle v, Wv\rangle$ is evaluated as

$$\nabla J(v) = Wv,$$

we can move v_k in the direction of negative gradient $-Wv_k$ to the global minimum point v^*. Systematically, the term Δv_k can be defined as follows:

$$g_k \triangleq Wv_k, \tag{4.1}$$

$$\Delta\widetilde{v_k} \triangleq -\omega \cdot g_k, \tag{4.2}$$

$$\widetilde{v_k} \triangleq \frac{v_k + \Delta\widetilde{v_k}}{\|v_k + \Delta\widetilde{v_k}\|}, \tag{4.3}$$

$$\Delta v_k \triangleq \widetilde{v_k} - v_k, \tag{4.4}$$

where ω is a small real factor to keep $\Delta\widetilde{v_k}$ and Δv_k small perturbations, and the formation (4.1)–(4.4) ensures $\|v_k + \Delta v_k\| = 1$, which is required in the variant of problem and perturbation analysis. Following the preceding procedure, we let

$$v_{k+1} = v_k + \Delta v_k, \tag{4.5}$$

thus constructing a series of v_k, whose convergence can be shown by the following proposition.

PROPOSITION 4.1. *Under the definition* (4.1)–(4.5), *let* v^* *be the unit eigenvector associated with the smallest eigenvalue of* W, *denoted as* $\lambda_{\min}(W)$, *which is single. If the initial vector* v_0 *is nonorthogonal to* v^* *and the factor* ω *satisfies* $\omega < \frac{1}{\lambda_{\max}(W)}$, *then the series* $\{v_k\}$ *converges to* $\pm v^*$.

Proof. From (4.1)–(4.5), we have

$$\begin{aligned} v_k &= \frac{v_{k-1} - \omega W v_{k-1}}{\|v_{k-1} - \omega W v_{k-1}\|} \\ &= \frac{(I - \omega W)v_{k-1}}{\|(I - \omega W)v_{k-1}\|} \\ &\triangleq \frac{(I - \omega W)v_{k-1}}{\alpha_{k-1}}. \end{aligned}$$

By recurrence, we have

$$v_k = \frac{(I - \omega W)^2 v_{k-2}}{\alpha_{k-1}\alpha_{k-2}} = \cdots = \frac{(I - \omega W)^k v_0}{\alpha_{k-1}\alpha_{k-2}\cdots\alpha_0}.$$

As the unit of $\|v_k\|$, it is obvious that

$$v_k = \frac{(I - \omega W)^k v_0}{\|(I - \omega W)^k v_0\|}.$$

When $0 < \omega < \frac{1}{\lambda_{\max}(W)}$, $\lambda(I - \omega W) = I - \omega\lambda(W) > 0$. In this case, the principal eigenvalue of $I - \omega W$ is $I - \omega\lambda_{\min}(W)$ and its associated eigenvector is $\pm v^*$. As the theory of power method, $(I - \omega W)^k v_0 \to \beta v^*$, where β is some real. So $v_k \to \pm v^*$. □

It is seen from Proposition 4.1 that the smallness of ω not only restricts the modulus of v_k but also ensures the convergence of the vector series. In the preceding discussion, we optimize $J(v)$ with respect to variable v along the gradient $g = Wv$, and ω plays the role of step length. It should be noticed that setting ω as a certain number less than $\frac{1}{\lambda_{\max}(W)}$ is sufficient for the convergence. It is different from the gradient method described in [7], where $J(b)$ is optimized along the gradient $\frac{DJ(b)}{\delta b}$,

ALGORITHM 1. Basic algorithm for minimization of $J_i(b)$.

Require: initial design b_0 ($b_0 \in \mathcal{A}$);
Ensure: optimized design b^*
1. **for** $k = 0$ **to** K **do**
2. **if** $k = 0$ **then**
3. compute the eigenvector v_k ($\|v_k\| = 1$); associated with the ith eigenvalue of B_kAB_k;
4. **else**
5. find $v_k = \text{argmin}\{\|v_k - v_{k-1}\| : v_k \text{ is an eigenvector of } B_kAB_k, \|v_k\| = 1\}$;
6. **end if**
7. check for convergence of v_k or $J_i(b_k)$;
8. **if** v_k or $J_i(b_k)$ converges **then**
9. **return** b_k;
10. **end if**
11. compute Δv_k;
12. compute Δb_k, i.e. solve equations $(\lambda_k I + B_kAB_k)B_k^{-1}\text{diag}(v)\Delta b_k = -(B_kAB_k - \lambda_k I)\Delta v_k$;
13. $b_{k+1} \leftarrow P(b_k + \Delta b_k)$
14. **end for**

and a line search subroutine is needed to determine the appropriate step length in each iteration.

After getting b_k, we are to compute the eigenvector v_k of B_kAB_k, which is continuously from the eigenvector v_{k-1} of $B_{k-1}AB_{k-1}$, assumed to belong to the ith eigenvalue. As the possible eigenvalues cross, it cannot be expected that v_k still belongs to the ith eigenvalue of B_kAB_k. The approach to locate v_k among all the eigenvectors is similar to that in [7]. That is, if Δb_{k-1} is small enough, all the eigenvectors of B_kAB_k come from their previous eigenvectors of $B_{k-1}AB_{k-1}$ continuously and slightly. As the symmetry of B_kAB_k, v_k is orthogonal to any other eigenvector, denoted as v_k'. Therefore, v_{k-1} must be nearly orthogonal to v_k', i.e., $\|v_{k-1} - v_k'\| \approx \sqrt{2} \gg \|v_{k-1} - v_k\|$. So we can just choose the one closest to v_{k-1} among all the unit eigenvectors of B_kAB_k to be v_k. In summary, the algorithm is shown in Algorithm 1.

The operator P in line 13 of Algorithm 1 means restricting $b_k + \Delta b_k$ to the admissible set $\mathcal{A}$. Considering that the objective $J_i(b_k)$ remains unchanged after multiplying a nonzero real number to b_k, it is hoped we find a real μ_k to make $\mu_k b_k$ as close as possible to $P(\mu_k b_k)$, thus weakening the restriction. One of the choices of P is as follows. Define

$$\widetilde{P}(b)^{(i)} \triangleq \begin{cases} b^{(i)}, & b_0 \leqslant b^{(i)} \leqslant b_1 \\ b_0, & b^{(i)} < b_0 \\ b_1, & b^{(i)} > b_1 \end{cases} \quad \text{for } i = 1, 2, \ldots, n,$$

and $\mu \triangleq \text{argmin}\{\theta(\mu b, \widetilde{P}(\mu b)) : \mu \in \mathbb{R}\}$, where $\theta(\cdot, \cdot)$ means the included angle between two vectors, and let $P(b) \triangleq \widetilde{P}(\mu b)$.

For line 5, we give a simple approach. In $BABv = \lambda v$, by perturbing b with Δb and causing linearized response $\Delta\lambda$ in λ and Δv in v, we got the linearized result

$$\Delta\lambda = 2\langle v, BA\Delta Bv\rangle.$$

If $b_k + \Delta b_k$ is not changed much after being operated by P in line 13, i.e., $b_{k+1} \approx b_k + \Delta b_k$, then Δv_k can be seen as an approximation to $v_{k+1} - v_k$, thereby $\Delta\lambda_k \triangleq 2$

$\langle v_k, B_k A \Delta B_k v_k \rangle$ can be seen as an approximation to $\lambda_{k+1} - \lambda_k$. So after getting b_{k+1}, we can just compute the eigenpairs of $B_{k+1}AB_{k+1}$ near the point $\widetilde{\lambda}_{k+1} \triangleq \lambda_k + \Delta\lambda_k$ and then search for them for v_{k+1}, so line 5 can be implemented as follows:

i. $\Delta\lambda_{k-1} \leftarrow 2\langle v_{k-1}, B_{k-1}A\Delta B_{k-1}v_{k-1}\rangle$;

ii. $\widetilde{\lambda}_k \leftarrow \lambda_{k-1} + \Delta\lambda_{k-1}$;

iii. find $v_k = \operatorname{argmin}\{\|v_k - v_{k-1}\| : v_k$ is a eigenvector of B_kAB_k associated with the eigenvalue near $\widetilde{\lambda}_k$, $\|v_k\| = 1\}$.

Although $b_k + \Delta b_k$ may be not changed much (or not changed at all) by P in line 13 at the first several iterations, it may have more and more components outweighing the given bounds b_0 and b_1 as k increases, which will be restricted forcibly by P in the end of every iteration. Consequently, there is such a phenomenon that even though $b_k + \Delta b_k$ is more optimized than the preceding design b_k, $P(b_k + \Delta b_k)$ may be less optimized than b_k, resulting in $J(b_{k+1}) > J(b_k)$. To ensure $J(b_k)$ is nonincreasing, we simply add a line search procedure to Algorithm 1, by seeing Δb_k as a direction and optimizing b_k along it. Notice the relation

$$(\lambda_k I + B_k A B_k) B_k^{-1} \operatorname{diag}(v) \Delta b_k = -(B_k A B_k - \lambda_k I)\Delta v_k,$$

where $\Delta v_k \approx -\omega W v_k$; applying a step length to Δb_k is almost equivalent to applying a step length to ω. To sum up, we give the modified Algorithm 2 with an unfixed ω.

It should be clarified that the unfixed ω in Algorithm 2 is different from that in the discussion in the beginning of section 4, where the ω does not have to be halved

ALGORITHM 2. Modified algorithm for minimization of $J_i(b)$.

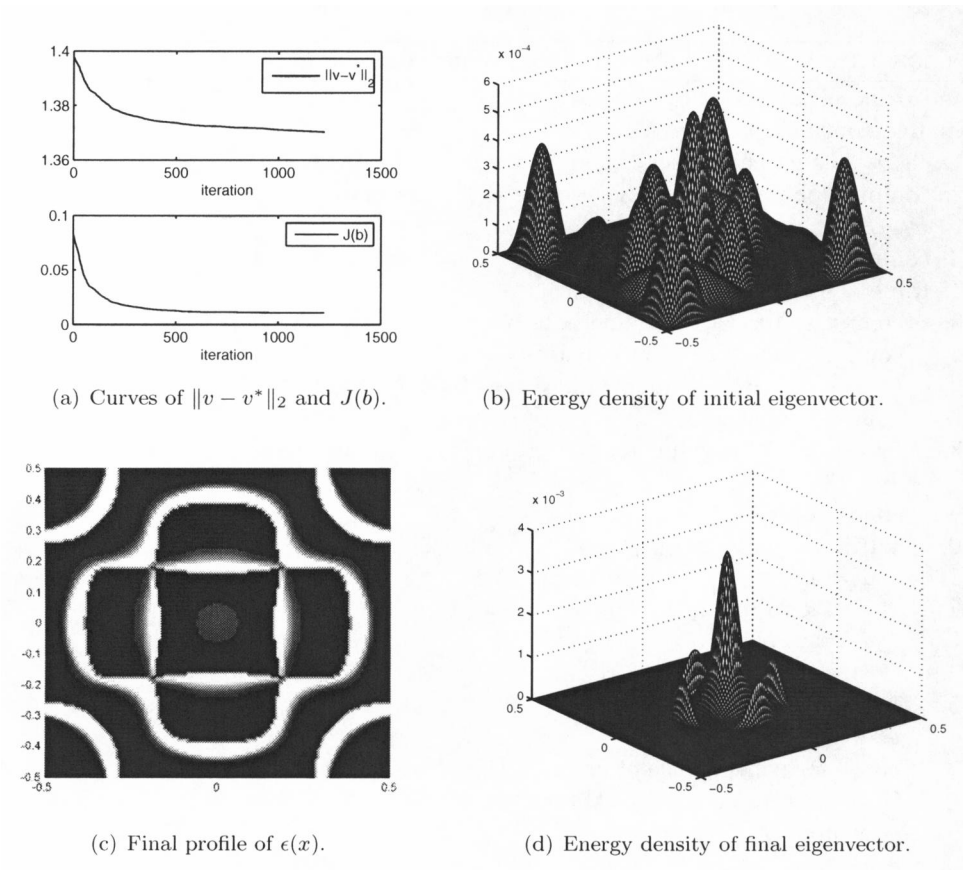
Require: initial design b_0 ($b_0 \in \mathcal{A}$);
Ensure: optimized design b^*

1. choose a step parameter ω and a tolerance *tol*;
2. compute the eigenvector v_0 ($\|v_0\| = 1$); associated with the ith eigenvalue of B_0AB_0;
3. compute $J_i(b_0)$;
4. **for** $k = 0$ **to** K **do**
5. compute Δv_k with parameter ω;
6. compute Δb_k, i.e. solve equations $(\lambda_k I + B_k A B_k) B_k^{-1} \operatorname{diag}(v) \Delta b_k = -(B_k A B_k - \lambda_k I)\Delta v_k$;
7. $b_{k+1} \leftarrow P(b_k + \Delta b_k)$
8. find $v_{k+1} = \operatorname{argmin}\{\|v_{k+1} - v_k\| : v_{k+1}$ is an eigenvector of $B_{k+1}AB_{k+1}$, $\|v_{k+1}\| = 1\}$;
9. compute $J_i(b_{k+1})$;
10. **while** $J_i(b_{k+1}) \geqslant J_i(b_k)$ **do**
11. $\omega = \omega/2$;
12. **if** $\omega < tol$ **then**
13. **return** b_k;
14. **end if**
15. same as Line 5-9;
16. **end while**
17. check for convergence of v_k or $J_i(b_k)$;
18. **if** v_k or $J_i(b_k)$ converges **then**
19. **return** b_k;
20. **end if**
21. **end for**

repeatedly to ensure the convergence. Here the ω is equivalent to being multiplied by a trial step. The only reason we add this step-halved line search subroutine is to overcome the interference of operator P, making $J(b_k)$ nonincreasing. All the following numerical experiments are based on Algorithm 2.

5. Numerical examples. We give three numerical examples to examine the algorithm. The domain Ω in problem (2.2) is set as $[-0.5, 0.5] \times [-0.5, 0.5]$ and is divided into a 120×120 grid. The Laplace operator is discretized by the five-point finite difference scheme. The weight matrix W is defined by $w(x) = |x|^2$, and the material constraints $\epsilon_1 \leftarrow 9$, $\epsilon_0 \leftarrow 1$. The algorithm is implemented through MATLAB, using its toolboxes and sparse matrix data structures if needed. In each iteration, the eigenpairs are computed by an implicitly restarted block Lanczos algorithm and the linear equations are solved by the generalized minimal residual method.

In the first example, the initial design $\epsilon(x) \equiv 25/16$ which is constant. We choose the 21st smallest eigenvalue, whose associated eigenvector has the energy distribution shown in Figure 5.1(b). We implement our method in this example. After nearly 1220 iterations, $J(b)$ falls from initial 8.218×10^{-2} to optimized 1.060×10^{-2} and converges. The curves of $\|v_k - v^*\|$ and $J(b_k)$, final distribution of $\epsilon(x)$, and energy density of final eigenvector v_f are shown in Figure 5.1(a), (c), (d).



(a) Curves of $\|v - v^*\|_2$ and $J(b)$.

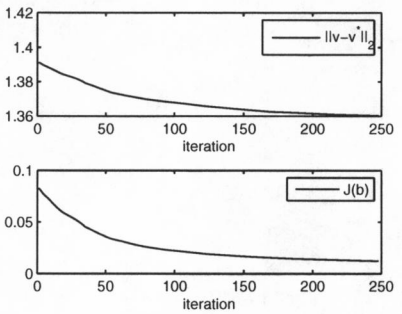
(b) Energy density of initial eigenvector.

(c) Final profile of $\epsilon(x)$.

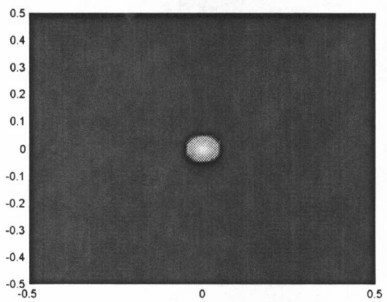
(d) Energy density of final eigenvector.

FIG. 5.1. *Variables for the first experiment.*

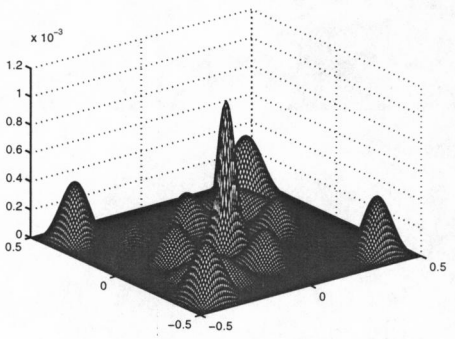
The second example is same as the first one, but the initial design $\epsilon(x) \equiv 25/16$ except for a small defect region in the center, as shown in Figure 5.2(b). The chosen eigenvector has the energy distribution shown in Figure 5.2(c). After 248 iterations, $J(b)$ falls from initial 8.273×10^{-2} to optimized 1.220×10^{-2} and a too-small ω stops the iteration. Compared with the result of the previous example, the final distribution of $\epsilon(x)$ in this example keeps its small defect region in the center. The curves of $\|v_k - v^*\|$ and $J(b_k)$, final distribution of $\epsilon(x)$, and energy density of final eigenvector v_f are shown in Figure 5.2(a), (e), (f).



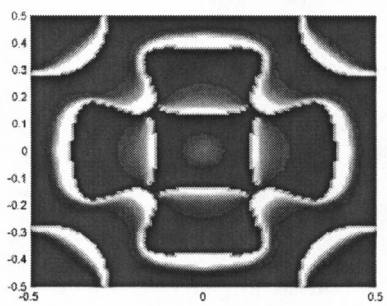
(a) Curves of $\|v - v^*\|_2$ and $J(b)$.



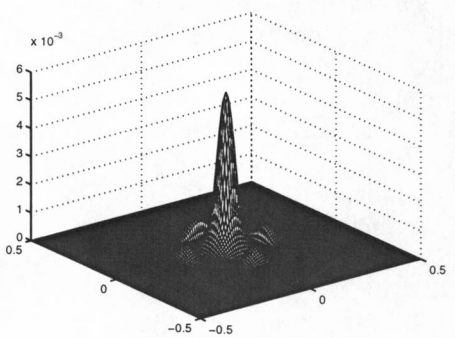
(b) Initial profile of $\epsilon(x)$.



(c) Energy density of initial eigenvector.

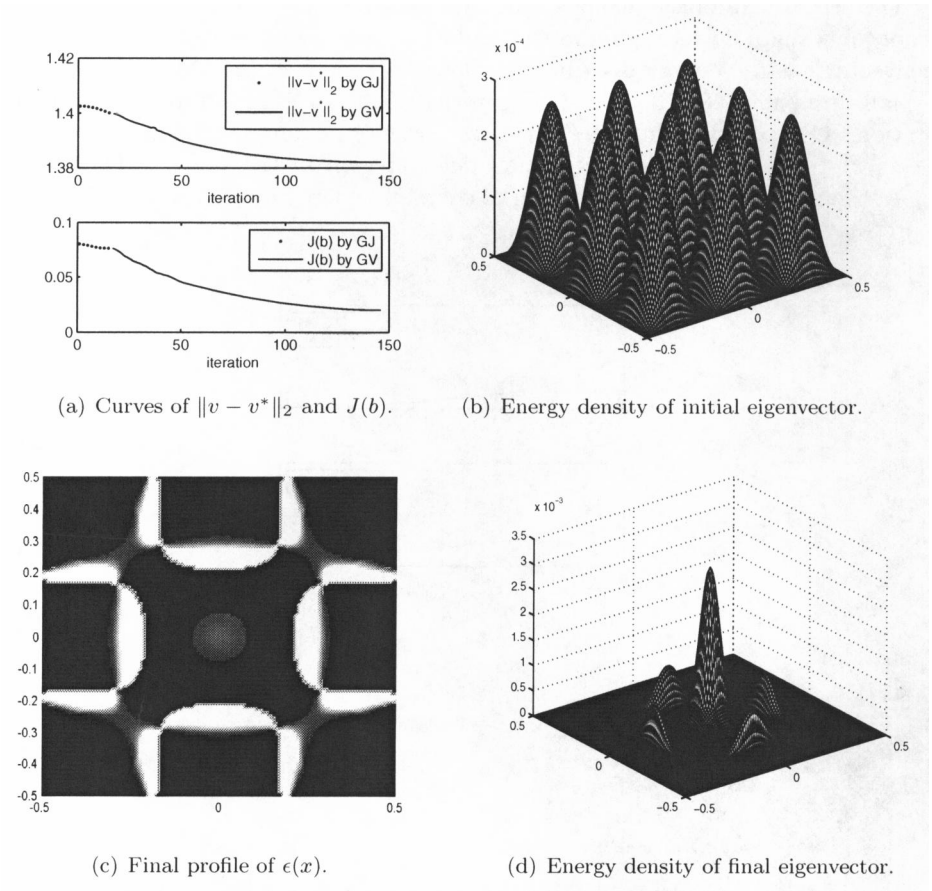


(d) Final profile of $\epsilon(x)$.



(e) Energy density of final eigenvector.

FIG. 5.2. *Variables for the second experiment.*



(a) Curves of $\|v - v^*\|_2$ and $J(b)$.

(b) Energy density of initial eigenvector.

(c) Final profile of $\epsilon(x)$.

(d) Energy density of final eigenvector.

FIG. 5.3. *Variables for the final experiment.*

In the final example, we keep track of the 11th smallest eigenvalue and set the initial design constant with value 25/16. The eigenvector has a lower frequency and degree of localization than in the previous examples, shown in Figure 5.3(b). First, we implement the gradient descent method (denoted GJ for its optimizing $J(b)$) described in [7]. The objective $J(b)$ falls from initial 8.03×10^{-2} to 7.63×10^{-2} after 16 iterations and gets stuck early. (It stops decreasing but increases in the next iteration.) Then we restart it by our algorithm (denoted GV for its optimizing $\|v - v^*\|$) and $J(b)$ decreases again to 2.00×10^{-2} after 136 iterations and converges. The curves of $\|v_k - v^*\|$ and $J(b_k)$ are shown in Figure 5.3(a), where the dotted line is produced by GJ and a solid line is produced by GV. The final distribution of $\epsilon(x)$ and the energy density of the final eigenvector are shown in Figure 5.3(c), (d). For comparison, we implement GV solely in this example, where it does not get stuck and gives a objective of 1.72×10^{-2}.

From the numerical results, we notice that our algorithm indeed works on the original optimization (2.7), though it is derived based on a variant (2.8). We can also notice that in most situations $J(b_k)$ and $\|v_k - v^*\|$ decrease or increase together, but not necessarily. In our examples, the phenomenon of eigenvalues crossing is observed. The tracked eigenvalue finally moves to the 36th, 36th, and 20th in the three examples

above. Furthermore, we should mention that the matrix M_k becomes more and more ill-conditioned during the iteration, because the eigenvector v_k goes close to $v^* = [0, \ldots, 0, 1, 0, \ldots, 0]$, leading more and more entries close to zero. Fortunately, from (3.7) we know the ill-conditioned factor contained in M_k is $\mathrm{diag}(v_k)$. Therefore it is not difficult to handle these ill-conditioned equations. For example, the entries of v_k which are less than a small given positive number can be set to zero, and the corresponding equations can be ignored. By solving the remaining system we will get a least square solution.

In the third example, the gradient descent method gets stuck early in 16th iteration, decreasing $J(b)$ only by nearly 5%. The reason, we believe, is the restriction of the bound operator P. In the description of the gradient method in [7], the operation of P is after the descent procedure, which may change the optimized design for the worse and increase the objective. This is verified in our extra tests by setting the bounds b_0 and b_1 to be more relaxed and observing the gradient method can work longer before getting stuck. However, our algorithm can get rid of the stuck point and restart this optimization in this case. Although the bound operator P also works in our algorithm, fortunately in most of our numerical examples the optimization can be done to a great extent with the modified Algorithm 2.

We should mention that the factor ω affects both the convergence and the convergence rate as we discusses in section 4. When ω is in a limited range, the series v_k will approach v^* more quickly as ω is set larger. But there may be no convergence if ω exceeds some threshold just as Proposition 4.1 shows. For another reason, we need to keep a small ω to guarantee the continuous variation of b.

In our numerical results, the phenomenon of a defect occurring in the center of the optimized medium is not obvious as in [7]. It should be pointed out that the process from an initial design to an optimized one through different methods may be different, because the current eigenvector v_k may approach the optimal one v^* in various directions, resulting in entirely different optimized designs. Recent related research is about design constraints and points out the optimal designs should be piecewise constant with upper and lower bound as values, namely, the "bang-bang" designs. For example, Osting [4] proves that the maximizer of the gap-to-midgap ratio of a special inhomogeneous wave equation is a bang-bang control, and the Bragg structure is the unique maximizer of the first spectral ratio.

6. Conclusion. This work studies an optimal design problem about photonic crystals, which requires increasing the localization of eigenfunctions of the Dirichlet eigenvalue problem. In the discrete form, we keep track of a certain eigenpair of the matrices formed by designs and make the objective as small as possible. Different from the gradient descent method that updates the variable b along the gradient of $J(b)$ in [7], we present a new method making the corresponding eigenvector v close to v^* along the gradient of $J(v)$. In the original steepest descent method, the gradient of $J(b)$ is computed directly from the eigenvector v, while in our method, the variation Δb is computed directly from Δv, which is chosen as a descent direction.

The numerical examples show the feasibility and convergence of our algorithm. On one side, it works well independently in some cases, where the objective is decreased below $1/6$ of its initial value. The final results obtained from a constant initial structure and a centered defect initial structure show a roughly consistent design but subtle differences. On the other side, our algorithm can restart the optimization if the gradient descent process stagnates at some stage because of the bound operator. It is shown that this algorithm works better than the gradient descent method in some

special cases, though some other approaches like disturbation may also be able to deal with the stagnation.

It should be noticed that one weakness of the algorithm derived in this paper is the successively singular coefficient matrices M_k, causing the solution of (3.5) away from the desired Δb, thus leading to a lower and lower convergence rate or even an extrema. Another weakness is the case-to-case property of this algorithm, which causes its limited generalization. After all, it applies to the specific objective (2.6), which can be intuitively linked to Rayleigh quotient property, then be transformed to the variant.

REFERENCES

[1] A. FIGOTIN AND A. KLEIN, *Localized classical waves created by defects*, J. Statist. Phys., 86 (1997), pp. 165–177.

[2] A. FIGOTIN AND A. KLEIN, *Midgap defect modes in dielectric and acoustic media*, SIAM J. Appl. Math., 58 (1998), pp. 1748–1773.

[3] B. OSTING, *Optimization of spectral functions of Dirichlet-Laplacian eigenvalues*, J. Comput. Phys., 229 (2010), pp. 8578–8590.

[4] B. OSTING, *Bragg structure and the first spectral gap*, Appl. Math. Lett., 25 (2012), pp. 1926–1930.

[5] C. Y. KAO, S. OSHER, AND E. YABLONOVITCH, *Maximizing band gaps in two-dimensional photonic crystals by using level set methods*, Appl. Phys. B, 81 (2005), pp. 235–244.

[6] C. Y. KAO AND F. SANTOSA, *Maximization of the quality factor of an optical resonator*, Wave Motion, 45 (2008), pp. 412–427.

[7] D. C. DOBSON AND F. SANTOSA, *Optimal localization of eigenfunctions in an inhomegeneous medium*, SIAM J. Appl. Math., 64 (2004), pp. 762–774.

[8] E. YABLONOVITCH, *Inhibited spontaneous emission in solid-state physics and electronics*, Phys. Rev. Lett., 58 (1987), pp. 2059–2062.

[9] I. M. KARABASH, *Optimization of quasi-normal eigenvalues for 1-D wave equations in inhomogeneous media: Description of optimal structures*, Asymptot. Anal., 81 (2013), pp. 273–295.

[10] P. HEIDER, D. BEREBICHEZ, R. V. KOHN, AND M. I. WEINSTEIN, *Optimization of scattering resonances*, Struct. Multidisciplinary Optim., 36 (2008), pp. 443–456.

[11] P. KUCHMENT, *The mathematics of photonic crystals*, in Mathematical Modeling in Optical Science, SIAM, Philadelphia, 2001, pp. 207–272.

[12] R. P. LIPTON, S. P. SHIPMAN, AND S. VENAKIDES, *Optimization of resonances in photonic crystal slabs*, in Physics, Theory, and Applications of Periodic Structures in Optics II, Proc. SPIE 5184, SPIE, Bellingham, WA, 2003, pp. 168–177.

[13] S. J. COX AND D. C. DOBSON, *Maximizing band gaps in two-dimensional photonic crystals*, SIAM J. Appl. Math., 59 (2004), pp. 2108–2120.

[14] S. J. COX AND D. C. DOBSON, *Band structure optimization of two-dimensional photonic crystals in H-polarization*, J. Comput. Phys., 158 (2000), pp. 214–224.

[15] S. JOHN, *Strong localization of photons in certain disordered dielectric superlattices*, Phys. Rev. Lett., 58 (1987), pp. 2486–2489.