

# Eigenvalue-accelerated LDOS optimization of high- $Q$ optical resonances

GEORGE SHAKER,<sup>1</sup> BEÑAT MARTINEZ DE AGUIRRE JOKISCH,<sup>2</sup>  
PENGNING CHAO,<sup>1</sup> AND STEVEN G. JOHNSON<sup>1,\*</sup>

<sup>1</sup>*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

<sup>2</sup>*NNF Quantum Computing Programme, Niels Bohr Institute, University of Copenhagen,  
Universitetsparken 5, 2100 Copenhagen, Denmark.*

\* [stevenj@math.mit.edu](mailto:stevenj@math.mit.edu)

**Abstract:** We demonstrate a new method that yields orders-of-magnitude acceleration in inverse design (e.g. topology optimization) of high- $Q$  resonant cavities to maximize the local density of states (LDOS), and which is also applicable to other resonant-response metrics. The key idea is that, once conventional LDOS optimization has identified a strong resonance, subsequent optimizations can exploit a fast shift-invert eigensolver to ensure that the LDOS remains centered at the resonance peak. We show that this eliminates ill-conditioning at sharp resonances that otherwise dramatically slows LDOS (and similar) optimization for  $Q \gg 100$ . Our method is demonstrated by design of  $Q > 10^6$  resonant cavities in 1d and 2d dielectric systems.

## 1. Introduction

Inverse design (large-scale optimization) [1, 2] of resonant effects in optics (or other wave systems) often proceeds by maximizing the frequency-domain response to sources at a given target frequency  $\omega_0$ , but severe challenges are known to arise for optimizing long-lifetime cavities. A resonant response that is often optimized is the power expended by a dipole source [3–8], which corresponds to the local density of states (LDOS, reviewed in Sec. 2), a key figure of merit for processes such as spontaneous emission [1, 2, 9, 10], and which is often approximated by the Purcell factor  $Q/V$  (the ratio of the mode quality factor  $Q$  to an appropriate measure of modal volume  $V$ ) [9]. That is, one maximizes  $\text{LDOS}(\omega_0, \mathbf{x}_0, \mathbf{p})$  at a fixed frequency  $\omega_0$  and dipole position  $\mathbf{x}_0$  over some geometry and/or material parameters  $\mathbf{p}$ . (More generally, one can maximize some function of the field intensity excited by a source or incident wave to optimize other resonant processes [11–24], as discussed below.) Maximizing such a quantity not only captures many physical effects simultaneously (e.g. lifetime, coupling, and spatial localization), but it is also amenable to efficient gradient-based optimization algorithms via adjoint methods [1, 3, 25] because the frequency-domain response is an easily differentiable function of material and geometric parameters. Although the resulting structures are resonant cavities characterized by high- $Q$  resonant modes [3, 4, 6, 8, 23], the optimization process need never identify a particular resonant mode (eigensolution) to optimize, which eliminates the need for a good initial guess of the optimal structure and avoids potential non-differentiability of eigenvalues [3]. However, all such methods have been observed to suffer from a serious problem when the  $Q$  (resonant lifetime) becomes sufficiently large ( $Q \gg 100$ ): optimization convergence becomes slower and slower [3], often requiring many thousands of iterations (Maxwell solves) even for  $Q \sim 10^4$ , making  $Q \gtrsim 10^5$  costly to obtain by maximizing LDOS or similar. The essential reason for this is that any field-intensity objective becomes like the “edge of a knife” at high  $Q$  (Sec. 2.1): performance falls off rapidly for perturbations that shift the resonant frequency away from  $\omega_0$ , but changes slowly for perturbations that keep the frequency fixed [3]. Optimization along such a sharp ridge is known to have poor performance due to the ill-conditioning of the second-derivative (Hessian) matrix [26].

In this paper, we demonstrate *orders-of-magnitude* acceleration in topology optimization (TO) [2] of high- $Q$  optical resonant responses by a new “eigenfrequency-shifted” approach

(Sec. 3.2) that compensates for the ill-conditioning arising from frequency sensitivity of high- $Q$  cavities [3]. We analyze our performance by explicit computation of the Hessian for a log LDOS objective, and show both analytically (Sec. 2.1) and numerically (Sec. 5.3) that frequency shifts correspond to a large Hessian eigenvalue,  $\sim Q^2$ , that is eliminated by our frequency-shifted objective, explaining our improved optimization conditioning. Our algorithm itself does not require a Hessian, and is amenable to efficient adjoint computation of gradients for optimization (Appendix A). We demonstrate our algorithm with TO of log LDOS for 1d (Sec. 5) and 2d (Sec. 6) dielectric cavities, including the non-intuitive case of enhancing light intensity *outside* the design region, obtaining  $Q > 10^8$  in 1d and  $Q > 10^6$  in 2d (for a tiny  $1.5\lambda_0 \times 1.5\lambda_0$  cavity). We also demonstrate further improvements by a “successive enlargement” heuristic, in which we gradually increase the diameter of the design region (Sec. 5). In 1d, we observe a clear relationship between the diameter of the design region and the resulting  $Q$  or LDOS, which increase exponentially in half-wave steps (Sec. 5).

A key idea is that we can initialize our new “shifted” optimization with  $\lesssim 10^3$  steps of the original “unshifted” method, which allows inverse design to discover a strong ( $Q \gtrsim 100$ ) resonance frequency  $\omega_*$  near  $\omega_0$  for subsequent eigensolver-based frequency-shift compensation. We can then optimize a frequency-shifted log-LDOS objective (Sec. 3.2):

$$\max_{\mathbf{p}} \log \text{LDOS}(\text{Re} [\omega_*(\mathbf{p}, \omega_0)], \mathbf{x}_0, \mathbf{p}), \quad (1)$$

where  $\omega_*(\mathbf{p}, \omega_0)$  is the (complex) eigenfrequency closest to  $\omega_0$ , with a constraint to ensure that the eigenvalue doesn’t drift too far from  $\omega_0$  during optimization. (We observe the logarithm to further accelerate convergence somewhat, similar to the use of  $\text{LDOS}^{-1}$  in previous work [3]; some optimization algorithms prefer objective functions to be of order unity [27, 28].) Although our approach involves additional cost per  $\mathbf{p}$  compared to the original “unshifted” methods, in order to perform a few steps of an iterative shift-invert eigensolve [29], we show that this is more than compensated by the accelerated  $\mathbf{p}$  convergence. As further evidence of our eigenfrequency-shifted approach’s performance relative to standard benchmarks, we show that we can narrow the gap between achievable LDOS and the theoretical upper bound [7, 30] in lossy materials.

Our examples in this paper maximize LDOS, but a similar approach could be applied (Sec. 7) to maximizing other frequency-domain resonant responses. For example, the LDOS has been replaced by other metrics for designing other incoherent emission processes [11] such as light-emitting diodes [12–15], scintillation [16, 17], or Raman scattering [18]. Additional linear-response objectives have been used to optimize Smith–Purcell radiation [19] or lasing efficiency [20]. One can even cascade multiple linear responses to maximize nonlinear effects such as second-harmonic generation [21, 22] or Raman effects [23, 24]. In all such cases,  $\text{Re} [\omega_*(\mathbf{p}, \omega_0)]$  can again be employed to stay on the resonance peak and accelerate convergence after an unshifted initialization identifies a target resonance. While we *exploit* this eigenvalue, we believe that one should generally strive for a holistic physical objective like LDOS that captures as much as possible of the physics of a given problem, including both lifetime ( $Q$ ) and localization effects (as opposed to specialized cases where one targets  $Q$  without optimizing localization [31–33] or targets a localization metric without consideration for  $Q$  [34, 35]).

## 2. Local density of states (LDOS)

In this section, we review the formulation of the time-harmonic Maxwell equations, resonant modes, the quality factor  $Q$ , and the local density of states (LDOS). For illustration purposes, we consider the 2d Maxwell equations in the  $xy$  plane for an out-of-plane ( $E_z$ ) polarization, although the general principles of the LDOS also apply to the full 3d Maxwell equations (as well as to other wave equations).

For a target excitation frequency  $\omega_0 > 0$ , a permittivity distribution  $\varepsilon(\mathbf{p}, \mathbf{x})$  for  $\mathbf{x} = (x, y) \in \mathbb{R}^2$ , and a source  $J_z(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_0)$ , the governing physics on the electric field  $E_z(\mathbf{x})$  are given by

the inhomogeneous Helmholtz equation

$$-\left(\nabla^2 + \omega_0^2 \epsilon\right) E_z = i\omega_0 J_z, \quad (2)$$

with Sommerfeld radiation boundary conditions, for dimensionless/natural units  $\epsilon_0 = \mu_0 = c = 1$ . The corresponding vacuum ( $\epsilon = 1$ ) wavelength is  $\lambda_0 = 2\pi/\omega_0$ ; below, we express distance in units of  $\lambda_0$  (equivalently,  $\lambda_0 = 1$  and  $\omega_0 = 2\pi$ ). Numerically, we discretize the Helmholtz equation using finite differences on a uniform grid—a finite-difference frequency-domain (FDFD) method [36–38]—to obtain

$$\underbrace{-\left(\mathbf{L} + \omega_0^2 \mathbf{D}\right)}_{\mathbf{A}} \mathbf{e} = \mathbf{b}, \quad (3)$$

where  $\mathbf{L}$  is a matrix approximating the second derivative,  $\mathbf{D} = \text{diag}(\epsilon)$ , and  $\mathbf{b}$  is a one-hot vector corresponding to a point source at  $\mathbf{x}_0$ . To implement outgoing boundary conditions, we surround the design region with perfectly matched absorbing layers (PML) [39], which are accounted for in the matrix  $\mathbf{L}$  as described in Appendix B. Henceforth, we will denote the matrix in Eq. (3) by  $\mathbf{A}(\omega_0, \mathbf{x}_0, \mathbf{p}) = -(\mathbf{L} + \omega_0^2 \mathbf{D})$ , omitting the arguments where convenient. These matrices are sparse, so we solved the system  $\mathbf{A}\mathbf{e} = \mathbf{b}$  using a sparse-direct algorithm [40, 41].

Electromagnetic reciprocity implies that the Helmholtz operator  $\nabla^2 + \omega^2 \epsilon$  is symmetric for scalar  $\epsilon$ , even in the presence of loss or absorbing boundaries [42, 43]. It is convenient below to choose the discretization to obtain the same property  $\mathbf{A} = \mathbf{A}^T$  for the matrix  $A$ , so that we do not need to distinguish left and right eigenvectors of  $\mathbf{A}$ . Although the typical stretched-coordinate PML formulation breaks this symmetry, the symmetry of  $\mathbf{A}$  can be restored by a simple diagonal scaling as described in Appendix B.

From our discretization, one may relate the quality factor to the governing physics of the Helmholtz equation with the following generalized eigenvalue problem:

$$-\mathbf{L}\mathbf{e}^{(k)} = (\mathbf{A} + \omega_0^2 \mathbf{D})\mathbf{e}^{(k)} = \omega_k^2 \mathbf{D}\mathbf{e}^{(k)}. \quad (4)$$

The quality factor  $Q$  of the mode  $\mathbf{e}^{(k)}$  with eigenfrequency  $\omega_k$  is defined as [9, 44]

$$Q = -\frac{\text{Re}[\omega_k]}{2\text{Im}[\omega_k]}. \quad (5)$$

The LDOS at  $\mathbf{x}_0$  is the power expended by a time-harmonic point dipole current source, equal to  $-\frac{1}{2} \text{Re} \left[ \int \mathbf{E} \cdot \mathbf{J} d\Omega \right]$  in Maxwell's equations [9], which in the discretized equations simplifies to:

$$\text{LDOS}(\omega_0, \mathbf{x}_0, \mathbf{p}) \sim -\text{Im} [\mathbf{e}^\dagger \mathbf{b}], \quad (6)$$

where we have dropped a factor of  $1/2\omega_0$  for simplicity—typically, our objective is the dimensionless LDOS *enhancement* relative to some reference system (e.g. vacuum), in which case such scale factors cancel. Henceforth, we denote the right-hand side of this expression as the “LDOS”. It is nonnegative by conservation of energy, allowing us to equivalently work with  $\log \text{LDOS}$  or  $\text{LDOS}^{-1}$  [3]. The LDOS is computationally convenient for optimization because it can be computed from the solution of a single system  $\mathbf{A}\mathbf{e} = \mathbf{b}$  and is a smooth function of any smooth parameterization  $\mathbf{p}$  of the discretized  $\epsilon$ , easily differentiable by adjoint methods [1, 3, 25, 45] as reviewed in Appendix A.

In 3d, or in 2d with the  $H_z$  polarization, the LDOS can diverge at sharp corners [46, 47], so to obtain a finite optimum one would need to regularize the problem [48] by imposing a minimum lengthscale [6–8, 48], specifying a nonzero separation between the emitter and the material [7, 46], and/or incorporating nonlocal effects [49, 50]. However, these concerns are independent of the ill-conditioning arising from high  $Q$ , and such singularities do not occur in 2d for the  $E_z$  polarization, so we need not consider them here.

### 2.1. $Q$ dependence of LDOS second derivatives

Consider a high- $Q$  resonance at a complex frequency  $\omega_*(p)$  that depends on some parameter(s)  $p$  of the geometry. That is,  $Q = -\text{Re}[\omega_*]/2\text{Im}[\omega_*] \gg 1$ . At frequencies  $\omega$  near  $\omega_*$ , for sources that couple strongly to the resonant mode, the LDOS is dominated by the resonant response and takes the form of a sharp peak. In particular, it is well known that the LDOS near a single strong resonance is approximately a Lorentzian peak [9]:

$$\text{LDOS}(\omega, p) \approx \frac{A(p)\gamma(p)}{\delta(p, \omega)^2 + \gamma(p)^2}, \quad (7)$$

for linewidth  $\gamma = -\text{Im}[\omega_*(p)] = \text{Re}[\omega_*(p)]/2Q$  and detuning  $\delta = \omega - \text{Re}[\omega_*(p)]$ , with some coupling strength  $A(p)$  depending on the geometry (but not depending directly on  $Q$ ).

For optimization over  $p$ , a key quantity is the second derivative [26], which is very large near such a sharp resonant peak, since any change in  $p$  will tend to shift the resonance frequency  $\text{Re}\omega_*$  away from the target frequency  $\omega = \omega_0$ . This can be quantified using the Lorentzian model. One finds that the second derivative at  $\delta = 0$  (the peak) with respect to a scalar  $p$  is:

$$\left. \frac{\partial^2 \text{LDOS}}{\partial p^2} \right|_{\omega=\text{Re}\omega_*} \approx \frac{A''}{\gamma} - \frac{Ay'' + 2A'\gamma'}{\gamma^2} + \frac{2A(\gamma'^2 - \delta'^2)}{\gamma^3}, \quad (8)$$

where primes denote derivatives with respect to  $p$ , and the dominant  $-2A\delta'^2\gamma^{-3} = O(Q^3)$  term is precisely  $\frac{\partial^2 \text{LDOS}}{\partial \omega^2} \left( \frac{\partial \text{Re} \omega_*}{\partial p} \right)^2$ . Applying the same analysis to log LDOS yields a dominant term that scales as  $O(Q^2)$ , since one factor of  $\gamma$  from the numerator separates out. More important than the absolute magnitude of the second derivative is the *ratio* of the large second derivative for frequency shifts ( $\delta'$ ) to the small second derivative for improvements in the amplitude ( $A''$ ), and this ratio also scales as  $O(\gamma^{-2}) = O(Q^2)$ , regardless of whether one looks at LDOS or log LDOS or  $\text{LDOS}^{-1}$ . (It is such ratios that determine the conditioning and influence the convergence rate of the optimization problem [26].)

Hence, we expect that the dominant term in our log LDOS Hessian should scale as  $O(Q^2)$  and arise from the dependence of the resonant frequency on the parameters. We confirm this scaling numerically in Sec. 5.3 for a vector  $\mathbf{p}$  of parameters: the dominant eigenvalue of the Hessian (second-derivative) matrix indeed scales with  $Q^2$  and the corresponding eigenvector is nearly parallel to  $\nabla_{\mathbf{p}} \text{Re} \omega_*$  (the direction of maximal frequency shift), exactly as predicted by this simple analysis. Thus, it is crucial to correct for resonant-frequency shifts in order to obtain a well-conditioned optimization problem.

## 3. Resonance/LDOS optimization

### 3.1. (Old) Unshifted algorithm

As proposed by Ref. 3 and adopted by several subsequent works [4–8], one can design a resonant cavity by directly maximizing the LDOS (or some monotonic function thereof, perhaps with bandwidth regularization via added loss [3]), i.e.:

$$\max_{\mathbf{p}} \log \text{LDOS}(\omega_0, \mathbf{x}_0, \mathbf{p}), \quad (9)$$

where  $\mathbf{p}$  is some parametrization of the geometry/materials; for example, in topology optimization (reviewed in Sec. 4) essentially “every pixel”  $\varepsilon$  of the design is a degree of freedom. As we commented for Eq. (1), we take the logarithm of the LDOS (which is always positive) because we observe this to yield slightly faster convergence, similar to the  $\text{LDOS}^{-1}$  suggested in Ref. 3, perhaps in part because many optimization algorithms favor objective functions scaled to be of order unity [27, 28] (due to dimensional hyperparameter choices in the algorithms).

A key advantage of directly maximizing the LDOS is that it is a smooth function of the geometry, and does not require one to select a particular resonant mode to optimize. In practice, one finds that maximizing LDOS (or similar figures of merit) quickly creates a resonant mode with a moderately high  $Q \gg 10^2$ , but then converges more slowly as  $Q$  (and hence LDOS) gradually increases [3, 48]. Examples of this behavior are given in Sec. 5 and Sec. 6, below. In some cases, to help the LDOS optimization locate a resonant mode, it can be effective to broaden the resonances by adding artificial absorption  $\alpha$  to the problem, e.g. multiplying  $\varepsilon$  everywhere by  $1 + i\alpha$ , and then decreasing  $\alpha \rightarrow 0^+$  as the optimization progresses, or alternatively employing  $\alpha > 0$  to impose a bandwidth regularization [3]. The key fact that LDOS maximization can quickly identify/create a good-quality resonant mode near the target frequency, even starting from a poor initial guess (such as vacuum or a random structure) is what we exploit for our new “shifted” objective in the next section.

### 3.2. (New) Shifted algorithm

After a few steps of unshifted optimization, one typically finds that the source excites a dominant resonance with  $Q \gtrsim 10^2$ . This provides us with an unambiguous resonant eigenfrequency for subsequent optimization: the eigenfrequency closest to the target  $\omega_0$ . Let  $\omega_*(\mathbf{p}, \omega_0)$  denote the eigenfrequency closest to  $\omega_0$  and let  $\mathbf{e}_*$  be the associated eigenvector. Since our unshifted initialization starts with a resonance quite close to  $\omega_0$  already, one can quickly compute  $\omega_*$  in a few iterations of efficient methods like shift-invert Arnoldi [29]. As shown in Appendix A, this requires only one additional expensive sparse-matrix factorization compared to the unshifted LDOS, including the gradient computation, approximately doubling the cost of each optimization step. Naively, one could then directly minimize something like  $|\omega_*(\mathbf{p}, \omega_0) - \omega_0|^2$  to obtain a high- $Q$  resonance near  $\omega_0$ , but that ignores the spatial-localization metrics of objectives like the LDOS. A hybrid objective (1), computing the LDOS (or similar) linear-response objective at the eigenfrequency  $\text{Re } \omega_*$ , captures all of the relevant physics while improving conditioning (keeping the optimizer centered on the peak frequency) and forestalling the need to differentiate an eigenvector-dependent objective (Sec. 7). As we shall show numerically in Sec. 5.3, changing the objective function to be consistently centered on  $\text{Re } \omega_*$  eliminates the  $\mathcal{O}(Q^2)$  eigenvalue of the Hessian. Hence, the optimizer is able to take larger steps and converges more rapidly. Our resulting “shifted” optimization problem is

$$\begin{aligned} \max_{\mathbf{p}} \quad & \log \text{LDOS}(\text{Re}[\omega_*(\mathbf{p}, \omega_0)], \mathbf{x}_0, \mathbf{p}) \\ \text{s.t.} \quad & \text{Re}[\omega_*(\mathbf{p}, \omega_0)] \in \text{BW}(\omega_0). \end{aligned} \tag{10}$$

Here, the bandwidth (BW) constraint is meant to prevent the optimizer from allowing  $\omega_*$  to drift far from  $\omega_0$  and to eliminate any instances of “jumping” to another mode. It merely serves a stabilizing purpose for the first few thousand iterations, and we found that it is not usually active throughout the optimization. Through experimentation, we found an appropriate bandwidth constraint for the test problems in this paper was often to simply bound  $\text{Re } \omega_*$  below by  $\omega_0$  or a nearby frequency, i.e.  $\text{BW}(\omega_0) = [\omega_0, \infty)$ , although this may vary from problem to problem. We discuss the choice of constraint further in Sec. 7. Since  $\text{Re } \omega_*$  and  $\nabla_{\mathbf{p}} \text{Re } \omega_*$  are already computed as part of the log LDOS objective and its gradient (Appendix A), adding one or more constraints on  $\text{Re } \omega_*$  is computationally “free” in that it requires no additional Maxwell solves. Fig. 1 depicts a schematic summary of our procedure.

## 4. Topology optimization (TO): Review

Having discretized the Helmholtz Eq. (3), we will employ density-based topology optimization (TO) [2], in which the parameters  $\mathbf{p}$  describe an artificial “density”  $\rho \in [0, 1]^N$  (one per grid

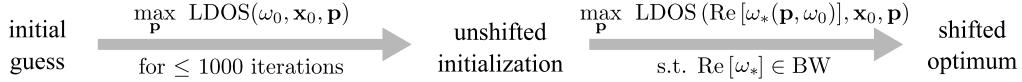


Fig. 1. Schematic of new eigenvalue-“shifted” algorithm: from an initial structure (e.g. vacuum) we run  $\lesssim 1000$  iterations of the original “unshifted” optimization of LDOS [3] to acquire an unshifted initialization which serves as the starting guess for the shifted optimization of Eqs. (1,10): maximizing the LDOS at the real part of nearest eigenfrequency  $\omega_*$ , constrained within some bandwidth BW, to ensure that the objective remains on the resonance peak.

point, in our case). This density can then be linearly mapped to the corresponding permittivity  $\varepsilon_k$  at each grid point  $k$  in the design region ( $\mathbf{x}_k \in \Omega$ ):

$$\varepsilon_k = \varepsilon_{\min} + (\varepsilon_{\max} - \varepsilon_{\min}) \rho_k , \quad (11)$$

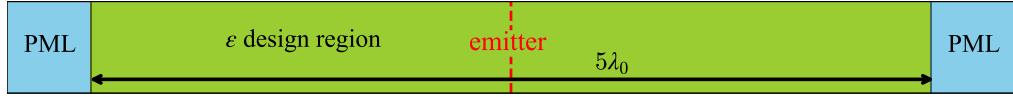
where  $\varepsilon_{\min} = 1$  and  $\varepsilon_{\max} = 12$  are the minimum and maximum permittivities, respectively. A key question is to precisely define how the parameters  $\mathbf{p}$  relate to the densities  $\rho$ . Typically, one takes  $\mathbf{p} \in [0, 1]^N$  to be density degrees of freedom, but does *not* directly set  $\rho = \mathbf{p}$ . Instead, one low-pass filters the  $\mathbf{p}$  in order to regularize the problem by imposing a minimum lengthscale, and then passes the resulting smoothed densities through some form of approximate Heaviside projection in order to approximately binarize the final density  $\rho$  to be nearly 0 or 1 almost everywhere [2].

In our 1d examples below, we omitted these filter/project steps and simply set  $\rho = \mathbf{p}$  (which is equivalent to directly optimizing over  $\varepsilon \in [\varepsilon_{\min}, \varepsilon_{\max}]$  at each grid point in the design region), similar to Ref. 3. It turned out that filtering was not necessary to regularize or project the 1d problem, since the designs in Sec. 5 invariably seemed to converge to something resembling a binary “quarter-wave stack” [44] where the layer thicknesses were on the order of a quarter wavelength in  $\varepsilon_{\min}$  or  $\varepsilon_{\max}$ . (Similar binary structures were observed in 2d, especially for the out-of-plane polarization, by Ref. 3, and there are theoretical arguments that unconstrained cavity optimization will generally lead to binary designs [51].)

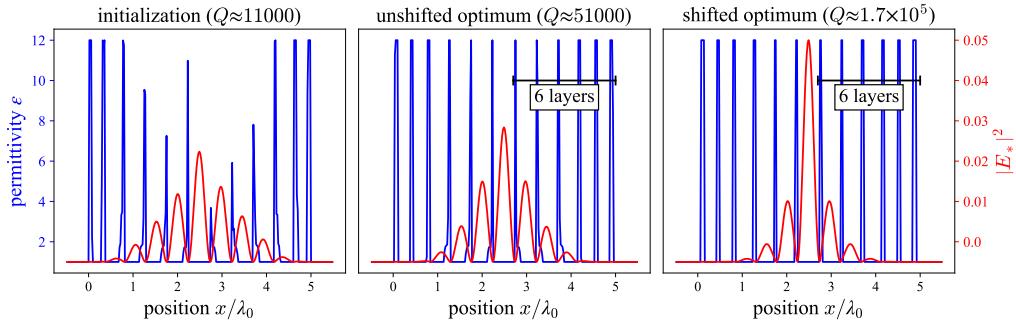
For our 2d results (Sec. 6), however, we found that unfiltered TO led to many pixel-scale features, and so we employed a filter-project algorithm using a conic filter [2] of radius  $0.1\lambda_0$ , along with a recently developed subpixel-smoothed projection (SSP) [52] algorithm. The projection (binarization) strength is controlled by a hyperparameter  $\beta$  that is essentially the steepness of a smoothed-Heaviside step function, which is gradually increased to binarize the structure during optimization while allowing the topology to smoothly change at early stages [2]. SSP allows us to set  $\beta$  to  $\infty$  in the final optimization stage in order to guarantee a structure that is binarized except in a 1-pixel layer at interfaces, while remaining differentiable (effectively a level-set method) [52]. In our 2d experiments below, we increased  $\beta$  through 8, 16, 40,  $\infty$  during the unshifted “initialization” stage to find a moderate- $Q$  resonance, but all subsequent 2d optimization was performed at  $\beta = \infty$  for fully binarized structures.

Given an implementation of an objective function such as  $f(\mathbf{p}) = \text{LDOS}(\mathbf{p})$ , here using the FDFD implementation from Sec. 2, one computes the gradient  $\nabla_{\mathbf{p}} f$  with respect to  $\mathbf{p}$ , at which point there is a wide selection of gradient-based optimization algorithms. The gradient is computed efficiently by backpropagating the derivatives through both the filtering and projection (via reverse-mode automatic differentiation [45, 53–55]) and through the PDE solver by manual implementation of an adjoint method [1, 25, 45] (described in Appendix A). We then employed an optimization algorithm by Svanberg *et al.* [27, 56] as implemented by the CCSAQ method of the free-software NLOpt package [57].

(a)



(b)



(c)

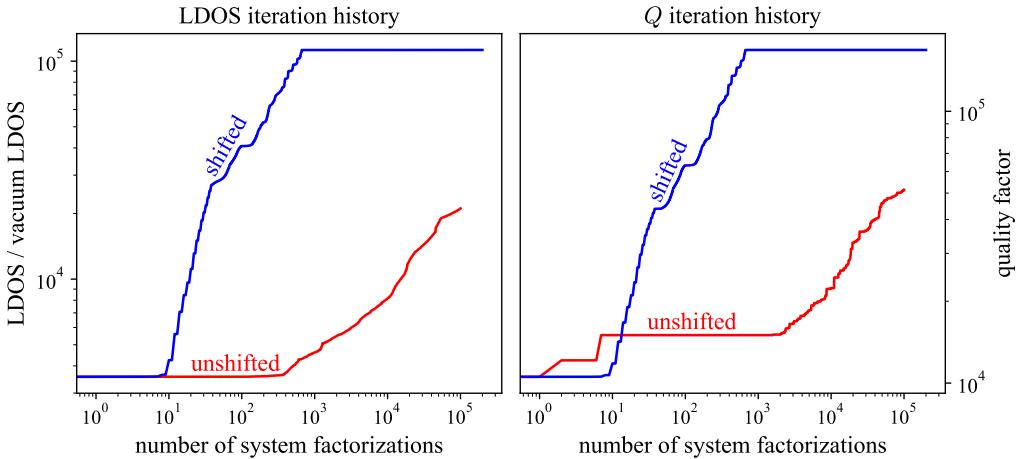


Fig. 2. (a) Schematic 1d optimization problem: maximizing LDOS for an emitter at the center for a design region that spans the whole computational cell except for terminating PML absorbing layers. (b) Unshifted initialization after 200 iterations (left) and shifted (right) and unshifted (middle) optima after  $10^5$  iterations: the resonant-mode profile (red, right axis) is superimposed over the permittivity  $\epsilon$  (blue, left axis). (c) LDOS enhancement (left) and resonance  $Q$  (right) as a function of the number of sparse-matrix factorizations (dominating the cost of the Maxwell solves) for the shifted (blue) and unshifted (red) algorithms.

## 5. One-dimensional results

To begin with, we consider a 1d problem in which the fields and materials only depend on  $x$ , with current  $J_z = \delta(x)$ . Such geometries allow rapid computational exploration, and are also easier to understand because the maximal localization in an infinite structure is expected to arise from the bandgap of a quarter-wave stack [44, 51], consisting of alternating layers of air (with thickness  $\lambda_0/4$ ) and dielectric (with thickness  $\lambda_0/4\sqrt{\varepsilon}$ , here with  $\varepsilon = 12$ ), at a given vacuum wavelength  $\lambda_0$ , within which a cavity is created by a “defect” (a perturbed layer thickness). So, one might expect similar alternating-layer structures to arise from inverse design, with LDOS and  $Q$  increasing exponentially with the diameter of the design region, albeit with some deviation from an exact quarter-wave stack due to the finite size. In 1d,  $Q > 10^8$  should be easily attainable with tractable computational domains. However, due to ill-conditioning, the convergence rate of inverse design will become very slow for high- $Q$  structures, and should be improved by our new eigenvalue-shifted algorithm.

### 5.1. Inverse design with a centered emitter

We consider a design region  $\Omega$  of width  $5\lambda_0$  (terminated on both sides by PML of thickness  $0.5\lambda_0$ ) with the source at the center of  $\Omega$ , as depicted in Fig. 2a. Our 1d FDFD grid has resolution  $50 \text{ pixels}/\lambda_0$ .

Starting from an initial  $\varepsilon = 1$ , we acquired an unshifted initialization, shown in Fig. 2b (left), by running 200 iterations of the unshifted algorithm. From this unshifted initialization, we then ran 100,000 iterations of the unshifted and shifted algorithms and found that the shifted algorithm quickly converges to a fully binary alternating-layer structure with  $Q \approx 1.7 \times 10^5$ , in Fig. 2b (right), while the unshifted algorithm stalls around  $Q \approx 5.1 \times 10^4$  (middle). The convergence history of the optimization algorithm, for both LDOS and  $Q$ , is shown in Fig. 2c, illustrating the dramatically better convergence of the shifted algorithm: after 100,000 iterations, the unshifted algorithm has improved LDOS by half as much as the shifted algorithm, which had already converged many tens of thousands of iterations ago, and the unshifted algorithm still has not converged. As described in Appendix A, the shifted algorithm requires twice as many sparse-matrix factorizations as the unshifted algorithm, approximately doubling the cost of each iteration, but we account for this by plotting the convergence as a function of the number of Maxwell-matrix factorizations. Both the unshifted and shifted optimizations are converging to cavities with 6 high-index layers on either side, identical to the initialization structure—we will see below that this is, in fact, only a local optimum, and a much better optimum can be found by a more careful initialization.

### 5.2. Successive enlargement

We find that even larger LDOS and  $Q$  can be obtained by a procedure of “successive refinement” of the degrees of freedom, a heuristic to accelerate convergence and/or evade poor local optima [58], often called “graduated” or “continuation” or “homotopy” optimization [59, 60]: gradually increasing the number of design degrees of freedom, using optimization results with fewer parameters as the starting guess for optimization with more parameters. In particular, we employ *successive enlargement* of the design domain, increasing the size  $L$  of the design region, using the optimum from the previous (smaller) design  $L_k$  as the successive starting guess for  $L_{k+1}$  as depicted in Fig. 3a. Our motivation for successive enlargement of the design domain in this case is evident from the structures of Fig. 2b: the optimization is clearly failing to prioritize the innermost layers of the cavity, which appear overly thin even for the shifted design, since the outer layers provide adequate confinement.

For comparison, we also considered a hand-designed quarter-wave stack [44] with a  $\lambda_0/2$  defect vacuum layer in the center, shown in Fig. 3c (left). It achieves  $Q \approx 4.0 \times 10^7$ , indicating that

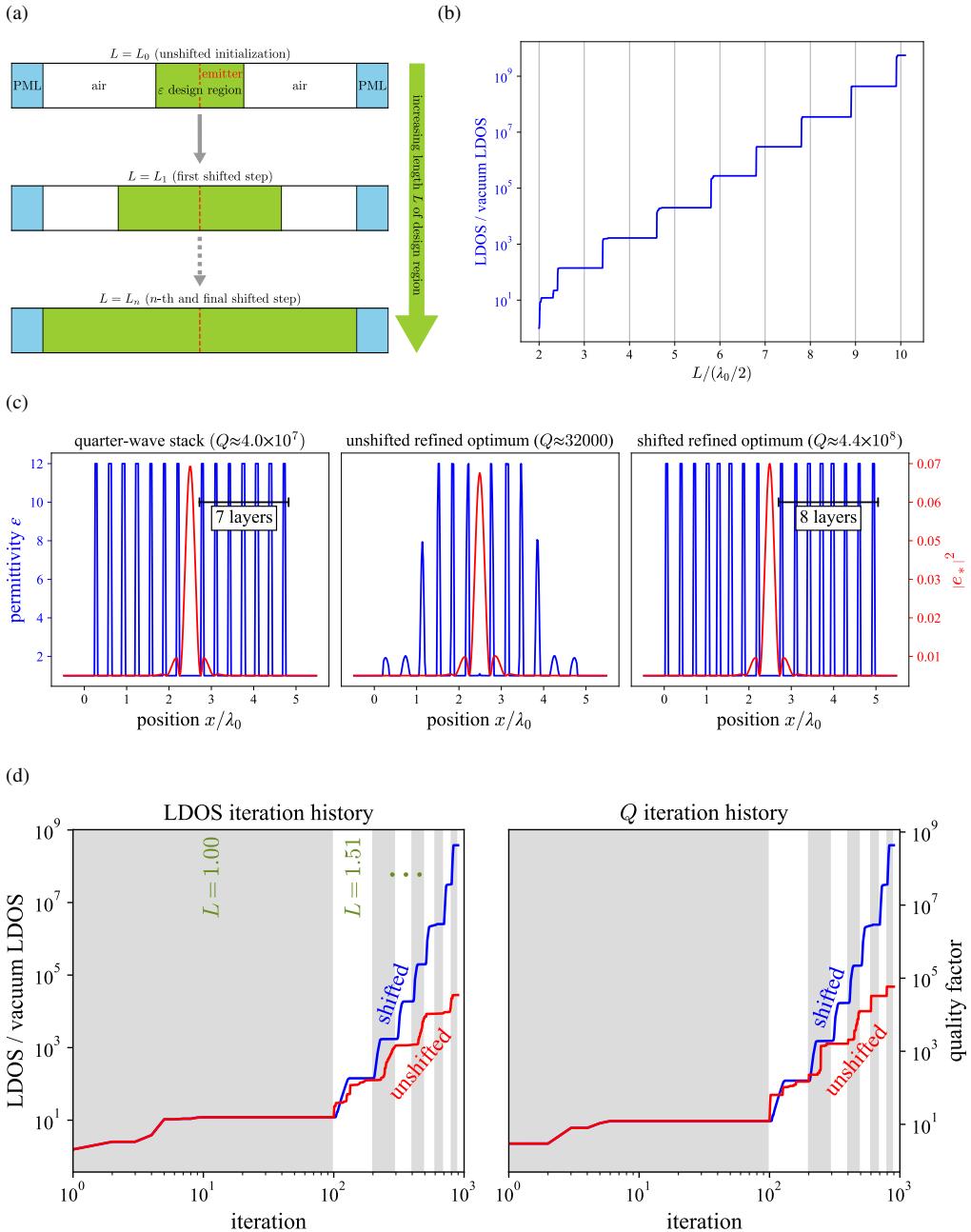


Fig. 3. (a) Schematic successive-enlarge algorithm in 1d: LDOS is optimized for a nested sequence of design domains  $L_0 < L_1 < \dots$  (green), which each optimum forming the initial guess for the next optimization. (b) LDOS enhancement as a function of design-region size  $L$ , exhibiting exponential growth in roughly half-wavelength steps. (c) A hand-designed quarter-wave-stack cavity (left), along with optimized cavities by successive enlargement with the unshifted (middle) and shifted (right) algorithms: the resonant-mode profile (red, right axis) is superimposed over the permittivity  $\epsilon$  (blue, left axis). (d) LDOS enhancement (left) and  $Q$  (right) as a function of the number of sparse-matrix factorizations for the successively-enlarged shifted (blue) and unshifted (red) algorithms. Alternating gray and white bands indicate a change in design-region size  $L$ .

even the shifted design from the previous section was suboptimal. Notice that this quarter-wave stack uses 7 high-index layers on either side of the cavity, suggesting that the previous 6-layer design is trapped in a different (topologically distinct) local optimum.

Using successive enlargement in conjunction with the shifted algorithm, however, corrects this deficiency of inverse design. The results are shown in Fig. 3, for which we have successively enlarged the design domain by increments of  $\Delta L = L_{k+1} - L_k = 0.51\lambda_0$ . The shifted algorithm converges to a cavity with  $Q = 4.4 \times 10^8$ , even better than the hand-designed cavity, in fewer than 1000 iterations, while the resulting structure in Fig. 3c (right) is more quarter-wave-like than before (with more uniform layer thicknesses). Moreover, it has 8 high-index layers on either side of the cavity, which is different from both the quarter-wave design and the previous non-successive optimization, suggesting the existence of multiple local optima. Successive enlargement with the unshifted algorithm yielded the same order of magnitude  $Q \sim 10^4$  as the previous optimization without refinement, although the resulting unshifted cavity structure in Fig. 3c (middle) now prioritizes the innermost layers, and the optimization history suggest that it has not yet fully converged. The convergence history, shown in Fig. 3d, illustrates rapid convergence of the shifted algorithm after each successive enlargement (alternating white/gray regions), whereas the unshifted algorithm is much slower. Here, for comparison purposes we employed 100 optimization iterations per design size  $L_k$ ; this is clearly more than was necessary for the shifted algorithm, while being insufficient for full convergence of the unshifted algorithm. If we had halted the shifted algorithm as soon as it converged for each  $L_k$ , many fewer ( $\ll 1000$ ) total iterations would be required for the same performance!

(One might wonder whether the successive optimization stages need to modify the innermost regions from the previous stages, or if it suffices to only optimize the *new* portions of the design domain. We attempted a “frozen-inner” variant of successive enlargement in which we kept the design in the inner  $L_k$  region fixed when optimizing the subsequent  $L_{k+1}$  domain, only allowing  $\varepsilon$  to vary in the added regions. We found that it was possible to get nearly as good LDOS and  $Q$ , within a factor of  $\approx 3$ , but that this result was extremely sensitive to the increment  $\Delta L$ .)

It is also interesting to investigate the dependence of the maximum LDOS (or the  $Q$ ) as a function of the design-region diameter  $L$ . To do this, we incremented  $L$  in very small steps  $L_{k+1} - L_k = 0.05\lambda_0$ , applying the shifted optimization algorithm, and plotted the resulting LDOS enhancement in Fig. 3b. What we observe is that there is a minimum increment  $\Delta L$  before LDOS can be substantially improved, corresponding to the minimum additional thickness before it is beneficial to begin to add another high-index layer. The theory of quarter-wave stacks suggests that this minimum  $\Delta L$  should be roughly  $\lambda_0/2$ , the thickness of two quarter-wave air layers (one on either side of the domain). Indeed, we observe that the increments  $\Delta L$  between successive “steps” in LDOS seem to be asymptotically approaching  $\approx \lambda_0/2$ , depicted by the gray grid lines in Fig. 3b. It would be good to investigate this phenomenon more rigorously by applying recent techniques to bound the LDOS for a given design region [7, 30], but this is outside the scope of the present work. (Ref. 51 also predicts an upper bound on  $Q(L)$ , but their bound on  $\log Q$  is proportional to  $L^2$  and appears to be extraordinarily loose, giving an upper bound  $Q \lesssim 10^{5000}$  for  $L = 5\lambda_0$  with  $\varepsilon \in [1, 12]$ .)

### 5.3. Conditioning of the Hessian

The 1d optimization problem is small enough that we can explicitly compute the Hessian (second-derivative) matrix, in order to verify the predictions of Sec. 2.1: the unshifted log LDOS objective should have a dominant eigenvalue scaling as  $O(Q^2)$ , with a corresponding eigenvector in the  $\text{Re}[\nabla_p \omega_*]$  direction corresponding to frequency shifts, while the shifted objective should remove this eigenvalue.

Computing a Hessian is well known to be costly: each column of the Hessian can be roughly as costly to compute as a single gradient computation [45, 54, 61]. In 1d with  $L = 5\lambda_0$ , we have

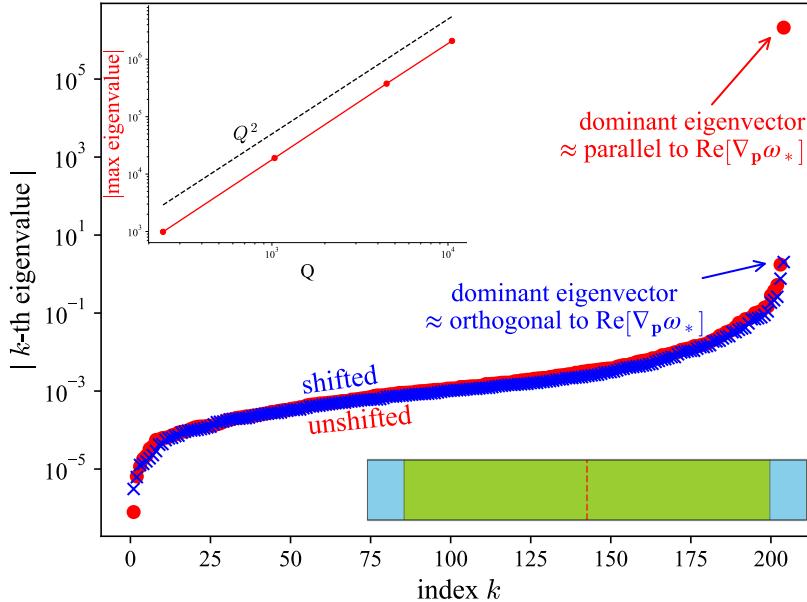


Fig. 4. Eigenvalues of the Hessians for the shifted (blue crosses) and unshifted (red dots) objective functions, computed at the unshifted initialization of our 1d test case (lower-right inset) from Fig. 2b(left). The largest eigenvalues differ by six orders of magnitude. Upper-left inset verifies that the dominant eigenvalue of the unshifted objective (red dots) scales like  $Q^2$  (dashed black line), as predicted in Sec. 2.1.

$n = 250$  parameters ( $50 \text{ pixels}/\lambda_0 \times 5\lambda_0$ ), requiring  $\gtrsim 250$  solves to obtain the entire Hessian even by the best known methods. Although there are sophisticated forward-over-reverse or reverse-over-reverse methods to compute the Hessian [45, 61], it was much easier to implement a simple finite-difference scheme that obtains the same scaling (but a worse constant factor). In particular, we applied a high-order finite-difference method, consisting of adaptive Richardson extrapolation of a centered difference [62, §5.7], to the gradient: each Hessian column is the partial derivative of the gradient with respect to one parameter. Once this is computed, we remove the rows and columns of the Hessian that correspond to parameters  $\mathbf{p}_k$  at their lower and upper bounds  $\{0, 1\}$ , since variables with active bound constraints are effectively removed from the optimization problem. The remaining  $\approx 200 \times 200$  Hessian is then diagonalized and we examine its eigenvalues and eigenvectors.

The result, applied to the unshifted initialization of Fig. 2b (left) with  $Q \approx 11000$ , is shown in Fig. 4. As expected, the Hessian of the unshifted log LDOS objective, whose eigenvalues are shown as red circles, is dominated by a single eigenvalue that is about  $10^5 \times$  the next-biggest eigenvalue. We find that the corresponding eigenvector is almost exactly parallel to  $\text{Re}[\nabla_{\mathbf{p}} \omega_*]$  (differing by  $< 10^{-4}$  degrees). To obtain the scaling with  $Q$ , shown in inset of Fig. 4, we applied the same procedure at 50, 100, 150, and 200 iterations of the unshifted algorithm starting from vacuum (yielding  $Q \approx 240, 1000, 4500, 11000$ , respectively); the resulting maximum Hessian eigenvalue scales almost exactly proportional to  $Q^2$ , as predicted. In contrast, when we compute the eigenvalues of the *shifted* log LDOS objective, shown as blue  $\times$ 's in Fig. 4, the largest eigenvalue of the unshifted Hessian is removed. The remaining dominant eigenvalue ( $\sim 10^5$  times smaller) corresponds to an eigenvector that is nearly *perpendicular* to  $\text{Re}[\nabla_{\mathbf{p}} \omega_*]$  (angle  $\approx 86^\circ$ ).

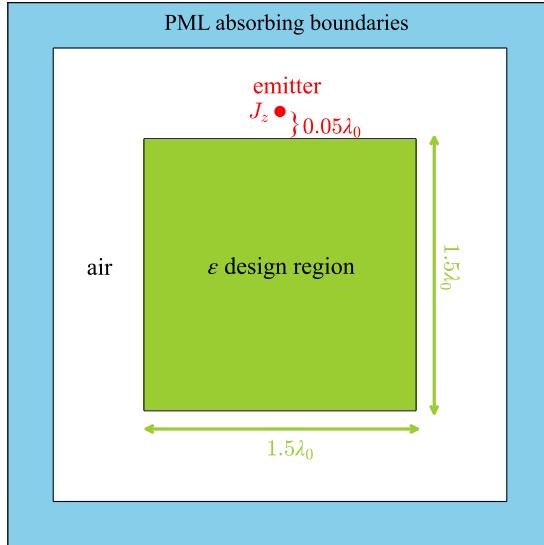


Fig. 5. Schematic example LDOS-optimization problem in 2d: a square  $1.5\lambda_0 \times 1.5\lambda_0$  design region (green), surrounded by air ( $\epsilon = 1$ ) and PML absorbers (blue) is optimized to maximize the LDOS of a point source (red dot)  $0.05\lambda_0$  outside the design region.

## 6. Two-dimensional results

In this section, we apply the unshifted and shifted optimization algorithms to a two-dimensional cavity, considering the challenging case of coupling to a point source that lies just *outside* the design region (so that a cavity cannot simply surround the source with a photonic-bandgap material [3, 44]). This scenario was also considered in recent theoretical upper bounds [7, 30].

In particular, Fig. 5. depicts our two-dimensional design region  $\Omega$ , a  $1.5\lambda_0 \times 1.5\lambda_0$  square region padded on all sides by  $0.5\lambda_0$  of air and surrounded by  $0.25\lambda_0$  of PML. The (out-of-plane) source is centered at a distance  $0.05\lambda_0$  above the upper edge of  $\Omega$ . The discretization has a resolution of  $20 \text{ pixels}/\lambda_0$ .

Starting from a uniform intermediate permittivity  $\epsilon = 6.5$ , we obtained the unshifted initialization shown in Fig. 6a (left), by running 1000 iterations of the unshifted algorithm. As explained in Sec. 4, we applied conic filter radius of  $0.1\lambda_0$  to set a minimum lengthscale in the design, and binarized with subpixel-smoothed projection at a sequence of increasing binarization strengths  $\beta$ : 100 iterations at  $\beta = 8$ , 200 at  $\beta = 16$ , 300 at  $\beta = 40$ , and 400 at  $\beta = \infty$ . From this initialization, both the shifted and unshifted algorithms were run for 500,000 iterations. The shifted algorithm converged to a design with  $Q \approx 1.6 \times 10^6$ , while the unshifted algorithm was still not fully converged at  $Q \approx 1.2 \times 10^5$ , as shown in Fig. 6a (middle, right). The convergence histories for both LDOS and  $Q$  are plotted in Fig. 6c, demonstrating that the shifted algorithm again exhibits substantially faster convergence and higher final performance. (As in Sec. 5, we plot convergence versus the number of Maxwell-matrix factorizations in order to account for the fact that the shifted iterations are roughly twice as expensive.) For example, the shifted algorithm reaches a  $Q$  of  $10^5$  in  $\approx 1000\times$  fewer iterations than the unshifted algorithm. The result resonant-mode  $|E_*|^2$  fields are shown Fig. 6b: they exhibit a tradeoff between strong spatial localization near the source and strong temporal confinement (high  $Q$ ).

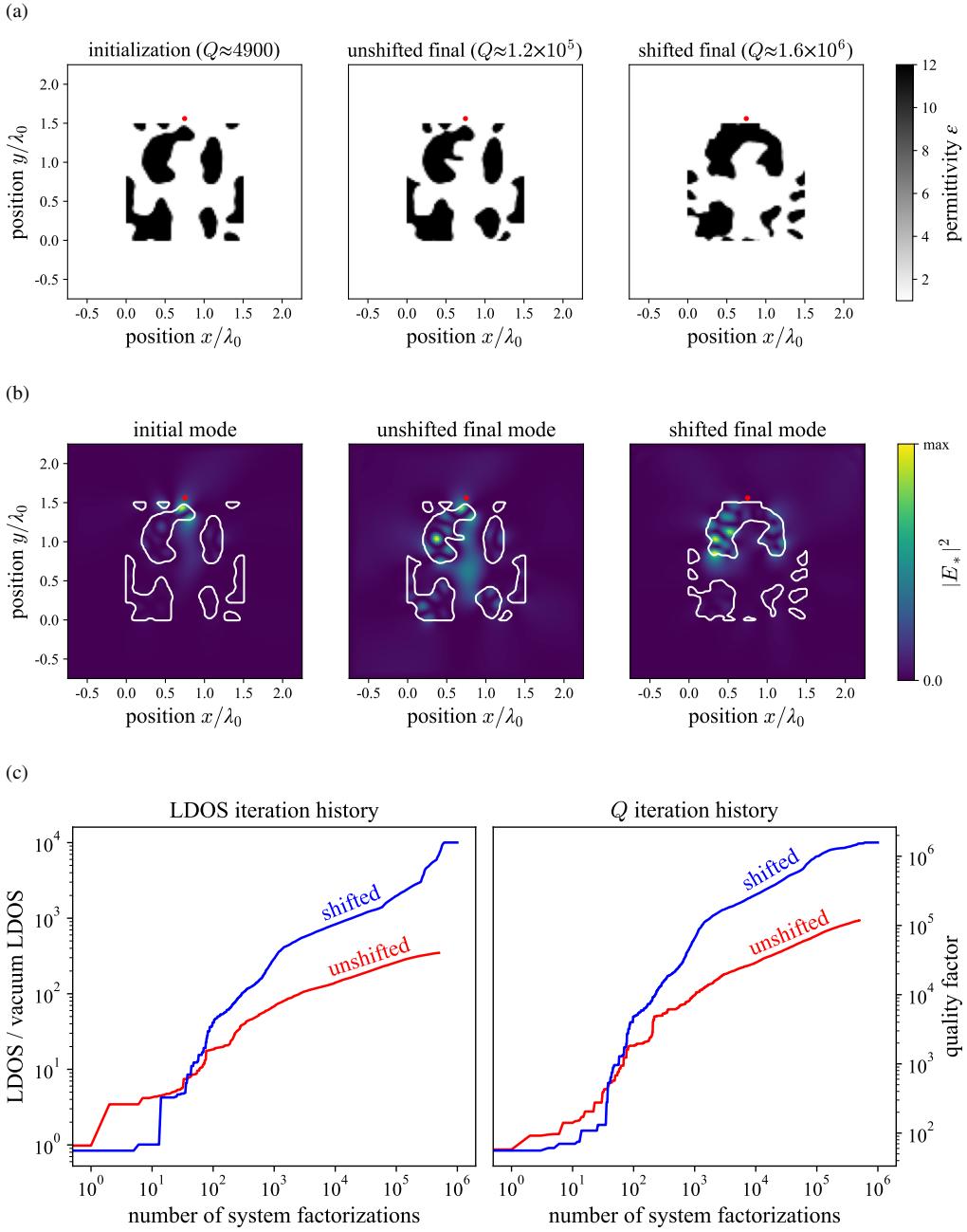


Fig. 6. (a) Unshifted initialization after 1000 iterations (left) and shifted (right) and unshifted (middle) final cavities after 500,000 iterations, with the source location (from Fig. 5) shown as a red dot. (b) Resonant-mode electric-field intensities for the initialization (left) and unshifted (middle) and shifted (right) final cavities, with the interface between air ( $\epsilon = 1$ ) and dielectric ( $\epsilon = 12$ ) superimposed in white. (c) LDOS enhancement (left) and  $Q$  (right) as a function of the number of sparse-matrix factorizations for the shifted (blue) and unshifted (red) algorithms.

### 6.1. Comparison to upper bounds

It is also interesting to compare to the theoretical upper bounds derived for a similar geometry [30]. These bounds were applied to a slightly different material  $\epsilon = 6 + 10^{-4}i$  that includes absorption loss ( $\omega \operatorname{Im} \epsilon > 0$ ), which helps limit the attainable  $Q$ . In Ref. 30, inverse design struggled to approach the theoretical upper bound, potentially due to a combination of slow convergence for high  $Q$  and being trapped in a poor local minimum: for an  $L \times L$  design region with  $L = 3\lambda_0$ , separated by the source at a distance  $0.2\lambda_0$ , unshifted inverse design from various random and deterministic initializations (including vacuum) with 15,000 iterations yielded an LDOS that was  $\approx 5\times$  below the theoretical bound. However, Ref. 30 showed that the upper-bound procedure could be exploited to predict a good initial guess for inverse design, from which unshifted inverse design yielded an LDOS  $\approx 1.6\times$  below the theoretical bound. Now that our shifted algorithm is not limited by slow convergence, we can apply it to the same problem and characterize the influence of the starting guess with greater certainty.

Applying our shifted optimization to the same  $3\lambda_0 \times 3\lambda_0$ ,  $0.2\lambda_0$  separation,  $\epsilon = 6 + 10^{-4}i$  design problem, with a simple *vacuum* initial structure followed by the same 1000-step unshifted initialization as above, we found that it converged within 20,000 shifted iterations. The resulting LDOS is about  $2.9\times$  below the theoretical bound. The corresponding lifetime is  $Q \approx 39000$ . This is significantly better than the previous unshifted inverse designs from similar starting points, suggesting that those previous results were limited by the slow convergence, although the  $Q$  in this case is small enough to practically reach with a larger number of iterations. However, it is still worse than optimization from a more sophisticated starting structure, suggesting a suboptimal local minimum.

## 7. Conclusion

Our results show that the new shifted LDOS-optimization algorithm can converge *orders of magnitude* faster than previous unshifted algorithms for  $Q \gg 1000$  structures. Like any nonconvex optimization, it can still obtain a suboptimal local minimum, but this problem can be ameliorated by improved starting structures obtained from theoretical bounds [30] or by heuristics such as successive enlargement of the design region. The drastic reduction in the number of optimization steps makes the shifted algorithm beneficial even though each evaluation of the objective function is more expensive (requiring an eigensolve), especially if the dominant cost is an explicit sparse-matrix factorization as in most previous resonance-optimization works: the shifted algorithm only doubles the number of factorizations compared to the unshifted method.

The same shifted algorithm should be applicable to other resonant-response objectives that have been optimized in previous work, such as nonlinear responses involving integrals of  $\|\mathbf{E}\|^4$  [18, 20] or compositions of multiple linear responses [21–24]. One should simply compute the *same* objective function at a different frequency given by the nearest resonant mode, after initializing with a few iterations of the unshifted algorithm to identify the resonance. There are also several other ways in which our shifted algorithm could potentially be improved. Rather than re-computing the LDOS or other objective at the shifted frequency, one could attempt to approximate the objective directly in terms of the computed resonant mode [9], but this would make gradient computation more complicated and costly: gradients of eigenvector-dependent objectives typically require the solution of an additional linear system [63, 64]. Another opportunity is to find better algorithms to prevent the eigensolver from “jumping” to a different nearby resonance, perhaps by monitoring changes in the eigenvector; currently, we have found that it suffices to impose a simple lower or upper bound on  $\operatorname{Re} \omega_*$ , but some experimentation may currently be required for each new system. Finally, it would also be valuable to devise algorithms that address *other* sources of slow convergence besides frequency shifts. For example, we found in 1d that successive enlargement greatly accelerated convergence in *addition* to escaping a poor local optimum, and it would be interesting to determine why this acceleration occurs and whether

related strategies can yield similar gains elsewhere.

## Appendix A: Gradient computation

In this section, we derive the gradients of the unshifted and shifted objective functions with respect to  $\boldsymbol{\varepsilon}$ . (These derivatives can then be propagated to the underlying “density” degrees of freedom via a straightforward chain rule, for which we applied automatic differentiation [53, 55].) The key goal is to obtain an “adjoint” (or “reverse-mode”) formulation in which computing the gradient has roughly the same cost as evaluating the objective function *once*, regardless of the number of parameters [1, 25, 45, 53, 54]. In fact, we find an even better result in both the unshifted and shifted cases: the gradient computation incurs *negligible additional cost* compared to the computation of the LDOS itself. To reduce the chance of an algebra error, all of our gradient formulas were validated numerically against a finite-difference approximation.

### 7.1. The unshifted gradient

First, let us review the gradient calculation for unshifted LDOS [3]. For all derivations that follow, let  $\mathbf{u}^{(k)}$  denote the  $k$ -th Cartesian basis vector; this is useful because our discretized Maxwell matrix  $\mathbf{A}$  only depends on  $\varepsilon_k$  (the permittivity at pixel  $k$ ) in the  $k$ -th diagonal element, so  $\frac{\partial \mathbf{A}}{\partial \varepsilon_k} = -\omega_0^2 \mathbf{u}^{(k)} (\mathbf{u}^{(k)})^T$  (a rank-1 matrix with only a single nonzero entry). We will also use the reciprocity identity [42, 43]  $\mathbf{A}^T = \mathbf{A}$  (enforced for our discretization by Appendix B). If  $\mathbf{e} = \mathbf{A}^{-1} \mathbf{b}$ , for a real vector  $\mathbf{b}$  ( $= \mathbf{u}^{(j)}$  at the source location  $j$ ), the chain rule then yields [3, 45]:

$$\begin{aligned} \frac{\partial \mathbf{b}^\dagger \mathbf{e}}{\partial \varepsilon_k} &= \mathbf{b}^T \frac{\partial \mathbf{A}^{-1}}{\partial \varepsilon_k} \mathbf{b} \\ &= -\mathbf{b}^T \mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \varepsilon_k} \mathbf{A}^{-1} \mathbf{b} \\ &= -\mathbf{b}^T \mathbf{A}^{-1} \left( -\omega_0^2 \mathbf{u}^{(k)} (\mathbf{u}^{(k)})^T \right) \mathbf{A}^{-1} \mathbf{b} \\ &= \omega_0^2 \mathbf{e}^T \mathbf{u}^{(k)} (\mathbf{u}^{(k)})^T \mathbf{e} \\ &= \omega_0^2 \mathbf{e}_k^2, \end{aligned}$$

where in the second-to-last line we employed reciprocity. Thus, since  $\text{LDOS} = -\text{Im}[\mathbf{e}^\dagger \mathbf{b}] = \text{Im}[\mathbf{b}^\dagger \mathbf{e}]$ , we have that

$$\frac{\partial \text{LDOS}}{\partial \varepsilon_k} = \omega_0^2 \text{Im} [\mathbf{e}_k^2], \quad (12)$$

which can be computed for all components  $k$  with *no* additional Maxwell solves beyond the solve for  $\mathbf{e}$  required to compute the LDOS itself: the gradient is essentially “free.”

### 7.2. The shifted gradient

We are now letting the frequency be a function of  $\mathbf{p}$  in our LDOS calculation:

$$\text{LDOS} := \text{LDOS}(\text{Re} [\omega_* (\mathbf{p}, \omega_0)], \mathbf{x}_0, \mathbf{p}). \quad (13)$$

This leads to an additional term in the chain rule, since the LDOS now depends on  $\mathbf{p}$  through both  $\varepsilon$  and  $\omega$ :

$$\frac{d\text{LDOS}}{d\mathbf{p}_k} = \left. \frac{\partial \text{LDOS}}{\partial \mathbf{p}_k} \right|_{\omega=\text{Re } \omega_*} + \left( \left. \frac{\partial \text{LDOS}}{\partial \omega} \right|_{\omega=\text{Re } \omega_*} \right) \text{Re} \left[ \frac{\partial \omega_*}{\partial \mathbf{p}_k} \right]. \quad (14)$$

The first term is given by Eq. 12. The partial derivative with respect to  $\omega$  can be computed as follows:

$$\begin{aligned}\frac{\partial \mathbf{b}^\dagger \mathbf{e}}{\partial \omega} &= \mathbf{b}^T \frac{\partial \mathbf{A}^{-1}}{\partial \omega} \mathbf{b} \\ &= -\mathbf{b}^T \mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \omega} \mathbf{A}^{-1} \mathbf{b} \\ &= -\mathbf{b}^T \mathbf{A}^{-1} (-2\omega \mathbf{D}) \mathbf{A}^{-1} \mathbf{b} \\ &= 2\omega \mathbf{e}^T \mathbf{D} \mathbf{e}.\end{aligned}$$

Thus,

$$\left. \frac{\partial \text{LDOS}}{\partial \omega} \right|_{\text{Re } \omega_*} = 2 \text{Re } \omega_* \cdot \text{Im} [\mathbf{e}^T \mathbf{D} \mathbf{e}], \quad (15)$$

which (as for  $\partial \text{LDOS} / \partial \varepsilon_k$  above) is essentially free: we re-use  $\mathbf{e}$  (from the Maxwell solve at the frequency  $\text{Re } \omega_*$ ), and  $\mathbf{D}$  is diagonal.

Now, let  $\mathbf{M} = \mathbf{D}^{-1} \mathbf{A}_0$ , where  $\mathbf{A}_0 = \mathbf{A}(\omega_0)$ , and let  $\mathbf{e}_*$  and  $\mathbf{f}_*$  be the right and left eigenvectors, respectively, of  $\mathbf{M}$  with eigenvalue  $\mu_* = \omega_*^2 - \omega_0^2$ . That is,  $\mathbf{M} \mathbf{e}_* = \mu_* \mathbf{e}_*$  and  $\mathbf{M}^\dagger \mathbf{f}_* = \overline{\mu_*} \mathbf{f}_*$ . From eigenvalue perturbation theory [65], we know that

$$\frac{\partial \mu_*}{\partial \varepsilon_k} = \frac{\partial \omega_*^2}{\partial \varepsilon_k} = \frac{(\mathbf{f}_*)^\dagger \frac{\partial \mathbf{M}}{\partial \varepsilon_k} \mathbf{e}_*}{(\mathbf{f}_*)^\dagger \mathbf{e}_*}. \quad (16)$$

(We computed  $\mathbf{e}_*$  using the shift-and-invert Arnoldi iteration [29, 66] applied to  $\mathbf{M}^{-1}$ , which requires only *one* additional sparse-matrix factorization of our Maxwell matrix  $\mathbf{A}_0$  at  $\omega_0$ .) By the symmetry of  $\mathbf{A}$  and  $\mathbf{D}$ , it follows that  $\overline{\mathbf{f}_*} = \mathbf{D} \mathbf{e}_*$ . In particular, combining this vector with the identity  $\mathbf{M}^T = \mathbf{D} \mathbf{M} \mathbf{D}^{-1}$  verifies the left eigen-equation:

$$\mathbf{M}^T \overline{\mathbf{f}_*} = \mathbf{D} \mathbf{M} \mathbf{D}^{-1} \overline{\mathbf{f}_*} = \mathbf{D} \mathbf{M} \mathbf{e}_* = \mu_* \mathbf{D} \mathbf{e}_* = \mu_* \overline{\mathbf{f}_*}. \quad (17)$$

We can now substitute into Eq. 16 to obtain

$$\frac{\partial \omega_*^2}{\partial \varepsilon_k} = \frac{\mathbf{e}_*^T \mathbf{D} \frac{\partial \mathbf{M}}{\partial \varepsilon_k} \mathbf{e}_*}{\mathbf{e}_*^T \mathbf{D} \mathbf{e}_*},$$

where

$$\begin{aligned}\frac{\partial \mathbf{M}}{\partial \varepsilon_k} &= \frac{\partial}{\partial \varepsilon_k} (\mathbf{D}^{-1} \mathbf{A}) \\ &= \frac{\partial}{\partial \varepsilon_k} (-\mathbf{D}^{-1} \mathbf{L} - \omega_0^2 \mathbf{I}) \\ &= -\frac{\partial}{\partial \varepsilon_k} (\mathbf{D}^{-1}) \mathbf{L} \\ &= \varepsilon_k^{-2} \mathbf{u}^{(k)} (\mathbf{u}^{(k)})^\dagger \mathbf{L}.\end{aligned}$$

Here,  $\mathbf{L}$  is the discretized Laplacian (from Sec. 2). Therefore, we have

$$\begin{aligned}\frac{\partial \omega_*^2}{\partial \varepsilon_k} &= \varepsilon_k^{-2} \frac{\mathbf{e}_*^T \mathbf{D} \mathbf{u}^{(k)} (\mathbf{u}^{(k)})^\dagger \mathbf{L} \mathbf{e}_*}{\mathbf{e}_*^T \mathbf{D} \mathbf{e}_*} \\ &= -\varepsilon_k^{-2} \omega_*^2 \frac{\mathbf{e}_*^T \mathbf{D} \mathbf{u}^{(k)} (\mathbf{u}^{(k)})^\dagger \mathbf{D} \mathbf{e}_*}{\mathbf{e}_*^T \mathbf{D} \mathbf{e}_*} \\ &= -\omega_*^2 \frac{(\mathbf{e}_*)_k^2}{\mathbf{e}_*^T \mathbf{D} \mathbf{e}_*}.\end{aligned}$$

We therefore obtain the last term required for the computation of the LDOS derivative above:

$$\frac{\partial \omega_*}{\partial \boldsymbol{\epsilon}_k} = -\frac{\omega_*}{2} \frac{(\mathbf{e}_*)_k^2}{\mathbf{e}_*^T \mathbf{D} \mathbf{e}_*}. \quad (18)$$

The computational cost of the shifted LDOS is roughly twice that of the unshifted LDOS, assuming that direct sparse factorization [40] of  $\mathbf{A}$  is employed, because the cost of this factorization typically dominates the computation time. That is, one now requires two factorizations: one at frequency  $\omega_0$  to use with the shift-invert eigensolver, and one at the new frequency  $\text{Re } \omega_*$ . However, given the shifted-LDOS computation, the gradient of the shifted LDOS is again essentially free, since it involves only simple computations with  $\mathbf{e}$  and  $\mathbf{e}_*$ , with a total computation time that scales linearly with the number of degrees of freedom (the number of pixels), much faster than sparse-matrix factorizations or solves.

## Appendix B: Preserving reciprocity $\mathbf{A} = \mathbf{A}^T$ with PML

Electromagnetic reciprocity implies that the linear operator relating fields and currents (or its inverse, the dyadic Green's function) is *complex-symmetric* assuming reciprocal materials  $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^T$  and  $\boldsymbol{\mu} = \boldsymbol{\mu}^T$  (but is not generally Hermitian in the presence of loss or outgoing boundary conditions) [42, 43]. For the algorithms in this paper, it is convenient to ensure that the corresponding matrix property  $\mathbf{A} = \mathbf{A}^T$  is preserved when the equations are discretized—this yields a simple identity between the left and right eigenvectors, both of which are needed for gradient computation (see Appendix A), so that they need not be computed separately. However, this symmetry depends on how one formulates the perfectly matched layer (PML) [39, 67] that implements the absorbing/outgoing boundaries. In this Appendix, we discuss how one can formulate the PML to preserve reciprocity.

PML can be elegantly derived from the “stretched-coordinate” viewpoint, in which one first analytically continues the equations to *complex* coordinates (in which waves become exponentially decaying without creating reflections), and then one performs a change of variables back to real coordinates [39, 67, 68]. For a PML along the  $x$  direction, this transforms every  $\frac{\partial}{\partial x}$  into  $\frac{1}{s_x(x)} \frac{\partial}{\partial x}$ , with a complex “stretch” factor is typically [39, 67, 69]  $s_x(x) = 1 + \frac{i\sigma_x(x)}{\omega}$ , where  $\sigma_x = 0$  in the interior of the domain (where solutions are unchanged) and  $\sigma_x/\omega > 0$  in the PML absorbing layers (inducing reflectionless attenuation) where  $\sigma_x$  ramps up gradually (often quadratically) to counteract discretization effects [39, 67, 69]. Similarly,  $\frac{\partial}{\partial y}$  is replaced by  $\frac{1}{s_y(x)} \frac{\partial}{\partial y}$ . In our 2d Helmholtz equation (2), this corresponds to the transformation

$$\nabla^2 + \omega^2 \boldsymbol{\varepsilon} \rightarrow \frac{1}{s_x} \frac{\partial}{\partial x} \frac{1}{s_x} \frac{\partial}{\partial x} + \frac{1}{s_y} \frac{\partial}{\partial y} \frac{1}{s_y} \frac{\partial}{\partial y} + \omega^2 \boldsymbol{\varepsilon}. \quad (19)$$

Unfortunately, this formulation breaks reciprocity (and leads to an asymmetric matrix  $\mathbf{A}$ ), because it has scale factors  $s_x^{-1}$  and  $s_y^{-1}$  on the left but not on the right. One can restore the symmetry with a diagonal similarity transformation  $\mathbf{A} \rightarrow \mathbf{S} \mathbf{A} \mathbf{S}^{-1}$  (a change of basis), however, because  $s_y$  commutes with  $\frac{\partial}{\partial x}$  and  $s_x$  commutes with  $\frac{\partial}{\partial y}$ . In particular, one can multiply on the left by  $\sqrt{s_x s_y}$  and divide on the right by  $\sqrt{s_x s_y}$ , yielding the symmetric operator

$$\frac{1}{\sqrt{s_x}} \frac{\partial}{\partial x} \frac{1}{s_x} \frac{\partial}{\partial x} \frac{1}{\sqrt{s_x}} + \frac{1}{\sqrt{s_y}} \frac{\partial}{\partial y} \frac{1}{s_y} \frac{\partial}{\partial y} \frac{1}{\sqrt{s_y}} + \omega^2 \boldsymbol{\varepsilon}. \quad (20)$$

This is the PML formulation used for the results in this paper, which we validated by the method of Ref. 69. Note that such transformations do not change the eigenvalues and merely scale the eigenvectors by  $\sqrt{s_x s_y}$  (as well as scaling the right-hand side currents by  $1/\sqrt{s_x s_y}$ ), which only changes the resonant mode profile and currents inside the PML (since  $\sqrt{s_x s_y} = 1$  in the interior

of the domain). Hence, it has no effect on figures of merit like LDOS that depend only on the interior fields (and use currents supported only in the interior), nor does it affect our sensitivity analysis of Appendix A. This transformation only changes the Laplacian matrix  $\mathbf{L}$  and not the diagonal material matrix  $\mathbf{D}$  in Eq. (3).

An alternative transformation that would accomplish the same goal of preserving reciprocity is to simply multiply both sides the field–current relation  $\mathbf{Ae} = \mathbf{b}$  or the eigen-equation  $\mathbf{Ae} = \mathbf{0}$  on the left by the diagonal operator (hence diagonal matrix)  $s_x s_y$ , yielding the operator

$$\frac{\partial}{\partial x} \frac{s_y}{s_x} \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \frac{s_x}{s_y} \frac{\partial}{\partial y} + \omega^2 s_x s_y \varepsilon, \quad (21)$$

which is also complex-symmetric, and again transforms the currents only in the PML regions (since  $s_x s_y = 1$  in the interior), but affects both  $\mathbf{L}$  and  $\mathbf{D}$ . (This is not technically a similarity transformation because we only multiplied on the left, but it does not change the matrix-pencil/generalized-eigenproblem solutions of  $\mathbf{Ae} = \mathbf{0}$ .) It is straightforward to derive that this corresponds exactly to the ‘‘UPML’’ formulation of PML [39, 69], in which the Jacobian scale factors of the stretched-coordinate formulation are converted into a modification of the fields and materials in the ordinary Maxwell equations by transformation optics [69]. In particular, the PML ‘‘stretch’’ factors correspond (in 3d) to a  $3 \times 3$  Jacobian matrix

$$\mathcal{J} = \begin{pmatrix} s_x^{-1} & & \\ & s_y^{-1} & \\ & & s_z^{-1} \end{pmatrix} \quad (22)$$

and the materials  $\varepsilon$  and  $\mu$  are transformed by [69]

$$\varepsilon \rightarrow \frac{\mathcal{J}\varepsilon\mathcal{J}^T}{\det \mathcal{J}}, \quad \mu \rightarrow \frac{\mathcal{J}\mu\mathcal{J}^T}{\det \mathcal{J}}, \quad (23)$$

which obviously preserves reciprocity  $\varepsilon = \varepsilon^T$  and  $\mu = \mu^T$ . Applied to the 3d Maxwell operator  $-\nabla \times \mu^{-1} \nabla \times + \omega^2 \varepsilon$  and simplified to the 2d case for  $z$ -invariant  $E_z$ -polarized fields with  $\mu = 1$  and scalar  $\varepsilon$ , this yields Eq. (21) above. (For iterative solvers, it has been shown that the stretched-coordinate formulation is more efficient than the UPML formulation [38], but one can always transform between formulations after the solver is complete. For sparse-direct solvers, all of these formulations, which differ only by diagonal scale factors, have similar costs.)

**Funding.** This work was supported in part by the U.S. Army Research Office through the Institute for Soldier Nanotechnologies (award no. W911NF-23-2-0121), by the Simons Foundation through the Simons Collaboration on Extreme Wave Phenomena, by the MIT Undergraduate Research Opportunities Program, and by the Novo Nordisk Foundation Quantum Computing Programme (award no. NNF22SA0081175).

**Disclosures.** The authors declare no conflicts of interest.

**Data availability.** Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

## References

1. S. Molesky, Z. Lin, A. Y. Piggott, *et al.*, ‘‘Inverse design in nanophotonics,’’ *Nat. Photonics* **12**, 659–670 (2018).
2. J. Jensen and O. Sigmund, ‘‘Topology optimization for nano-photonics,’’ *Laser & Photonics Rev.* **5**, 308–321 (2010).
3. X. Liang and S. G. Johnson, ‘‘Formulation for scalable optimization of microcavities via the frequency-averaged local density of states,’’ *Opt. Express* **21**, 30812 (2013).
4. F. Wang, R. E. Christiansen, Y. Yu, *et al.*, ‘‘Maximizing the quality factor to mode volume ratio for ultra-small photonic crystal cavities,’’ *Appl. Phys. Lett.* **113** (2018).
5. W. Yao, M. Benzaouia, O. D. Miller, and S. G. Johnson, ‘‘Approaching the upper limits of the local density of states via optimized metallic cavities,’’ *Opt. Express* **28**, 24185–24197 (2020). Editor’s Pick.

6. M. Albrechtsen, B. Vosoughi Lahijani, R. E. Christiansen, *et al.*, “Nanometer-scale photon confinement in topology-optimized dielectric cavities,” *Nat. Commun.* **13** (2022).
7. P. Chao, R. Kuaté Defo, S. Molesky, and A. Rodriguez, “Maximum electromagnetic local density of states via material structuring,” *Nanophotonics* **12**, 549–557 (2022).
8. G. İşiklar, P. T. Kristensen, J. Mørk, *et al.*, “On the trade-off between mode volume and quality factor in dielectric nanocavities optimized for purcell enhancement,” *Opt. Express* **30**, 47304–47314 (2022).
9. A. Oskooi and S. G. Johnson, “Electromagnetic wave source conditions,” in *Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology*, A. Tafove, A. Oskooi, and S. G. Johnson, eds. (Artech, Boston, 2013), chap. 4, pp. 65–100.
10. C. M. Cisowski, R. Kilianski, and R. Bennett, “Transforming photonics: Inverse design for optical cavity engineering,” *J. Opt.* **27**, 083002 (2025).
11. W. Yao, F. Verdugo, R. E. Christiansen, and S. G. Johnson, “Trace formulation for photonic inverse design with incoherent sources,” *Struct. Multidiscip. Optim.* **65**, 336 (2022).
12. O. T. A. Janssen, A. J. H. Wachters, and H. P. Urbach, “Efficient optimization method for the light extraction from periodically modulated leds using reciprocity,” *Opt. Express* **18**, 24522 (2010).
13. S. Kim, J. M. Shin, J. Lee, *et al.*, “Inverse design of organic light-emitting diode structure based on deep neural networks,” *Nanophotonics* **10**, 4533–4541 (2021).
14. A. Mayer and A. Bay, “Optimization by a genetic algorithm of the light-extraction efficiency of a GaN light-emitting diode,” *J. Opt.* **17**, 025002 (2015).
15. M. F. Schubert and A. M. Hammond, “Fourier modal method for inverse design of metasurface-enhanced micro-LEDs,” *Opt. Express* **31**, 42945 (2023).
16. C. Roques-Carmes, N. Rivera, A. Ghorashi, *et al.*, “A framework for scintillation in nanophotonics,” *Science* **375** (2022).
17. R. Schuetz, Y. Kurman, N. Lahav, *et al.*, “Purcell-enhanced X-ray imaging in ultra-thin scintillators,” in *CLEO 2023*, (Optica Publishing Group, 2023), CLEO\_AT, p. AW3Q.7.
18. W. Yao, F. Verdugo, H. O. Everitt, *et al.*, “Designing structures that maximize spatially averaged surface-enhanced Raman spectra,” *Opt. Express* **31**, 4964 (2023).
19. U. Haeusler, M. Seidl, P. Yousefi, and P. Hommelhoff, “Boosting the efficiency of Smith–Purcell radiators using nanophotonic inverse design,” *ACS Photonics* **9**, 664–671 (2022).
20. B. M. d. A. Jokisch, A. Cerjan, R. E. Christiansen, *et al.*, “Efficient first-principles inverse design of nanolasers,” *arXiv:2506.20223* (2025).
21. S. A. Mann, H. Goh, and A. Alù, “Inverse design of nonlinear polaritonic metasurfaces for second harmonic generation,” *ACS Photonics* (2023).
22. Z. Lin, X. Liang, M. Lončar, *et al.*, “Cavity-enhanced second harmonic generation via nonlinear-overlap optimization,” *Optica* **3**, 233–238 (2016).
23. R. E. Christiansen, J. Michon, M. Benzaouia, *et al.*, “Inverse design of nanoparticles for enhanced Raman scattering,” *Opt. Express* **28**, 4444 (2020).
24. Y. Pan, R. E. Christiansen, J. Michon, *et al.*, “Topology optimization of surface-enhanced Raman scattering substrates,” *Appl. Phys. Lett.* **119** (2021).
25. T. W. Hughes, M. Minkov, I. A. D. Williamson, and S. Fan, “Adjoint method and inverse design for nonlinear nanophotonic devices,” *ACS Photonics* **5**, 4781–4787 (2018).
26. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer New York, 2006).
27. K. Svanberg, “MMA and GCMMA: Two methods for nonlinear optimization,” <https://people.kth.se/~krille/mmagcmma.pdf> (2007).
28. A. Wächter and L. T. Biegler, “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” *Math. programming* **106**, 25–57 (2006).
29. Z. Bai, J. Demmel, J. Dongarra, *et al.*, eds., *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide* (SIAM, Philadelphia, 2000).
30. P. Chao, A. Amaolo, S. Molesky, and A. W. Rodriguez, “Bounds as blueprints: Towards optimal and accelerated photonic inverse design,” *arXiv:2504.10469* (2025).
31. W. R. Frei, H. T. Johnson, and K. D. Choquette, “Optimization of a single defect photonic crystal laser cavity,” *J. Appl. Phys.* **103** (2008).
32. G. H. Ahn, K. Y. Yang, R. Trivedi, *et al.*, “Photonic inverse design of on-chip microresonators,” *ACS Photonics* **9**, 1875–1881 (2022).
33. R. Li, C. Zhang, W. Xie, *et al.*, “Deep reinforcement learning empowers automated inverse design and optimization of photonic crystals for nanoscale laser cavities,” *Nanophotonics* **12**, 319–334 (2023).
34. J. Lu and J. Vuckovic, “Inverse design of nanophotonic structures using complementary convex optimization,” *Opt. Express* **18**, 3793 (2010).
35. R. E. Christiansen, J. Mørk, and O. Sigmund, “Orders-of-magnitude reduction in photonic mode volume by nano-sculpting,” *arXiv:2406.16461* (2024).
36. A. Christ and H. L. Hartnagel, “Three-dimensional finite-difference method for the analysis of microwave-device embedding,” *IEEE Trans. on Microw. Theory Tech.* **35**, 688–696 (1987).
37. R. C. Rumpf, *Electromagnetic and Photonic Simulation for the Beginner: Finite-Difference Frequency-Domain in*

- MATLAB* (Artech, Boston, MA, 2022).
38. W. Shin and S. Fan, “Choice of the perfectly matched layer boundary condition for frequency-domain Maxwell’s equations solvers,” *J. Comput. Phys.* **231**, 3406—3431 (2012).
  39. A. Tafove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Artech house, 2005).
  40. T. A. Davis, *Direct Methods for Sparse Linear Systems* (SIAM, Philadelphia, 2006).
  41. T. A. Davis, “Algorithm 832: UMFPACK v4.3—an unsymmetric-pattern multifrontal method,” *ACM Trans. on Math. Softw.* **30**, 196–199 (2004).
  42. W. C. Chew, “A new look at reciprocity and energy conservation theorems in electromagnetics,” *IEEE Trans. on Antennas Propag.* **56**, 970–975 (2008).
  43. C. Guo, Z. Zhao, and S. Fan, “Internal transformations and internal symmetries in linear photonic systems,” *Phys. Rev. A* **105** (2022).
  44. J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light* (Princeton University Press, 2008), 2nd ed.
  45. P. Bright, A. Edelman, and S. G. Johnson, “Matrix calculus (for machine learning and beyond),” [arXiv:2501.14787](https://arxiv.org/abs/2501.14787) (2025).
  46. O. D. Miller, A. G. Polimeridis, M. T. H. Reid, *et al.*, “Fundamental limits to optical response in absorptive systems,” *Opt. Express* **24**, 3329 (2016).
  47. H. Choi, M. Heuck, and D. Englund, “Self-similar nanocavity design with ultrasmall mode volume for single-photon nonlinearities,” *Phys. Rev. Lett.* **118** (2017).
  48. M. Chen, R. E. Christiansen, J. A. Fan, *et al.*, “Validation and characterization of algorithms and software for photonics inverse design,” *J. Opt. Soc. Am. B* **41**, A161–A176 (2024).
  49. C. Ciraci, R. T. Hill, J. J. Mock, *et al.*, “Probing the ultimate limits of plasmonic enhancement,” *Science* **337**, 1072–1074 (2012).
  50. G. Toscano, S. Raza, S. Xiao, *et al.*, “Surface-enhanced raman spectroscopy: nonlocal limitations,” *Opt. Lett.* **37**, 2538 (2012).
  51. B. Osting and M. I. Weinstein, “Long-lived scattering resonances and Bragg structures,” *SIAM J. on Appl. Math.* **73**, 827–852 (2013).
  52. A. M. Hammond, A. Oskooi, I. M. Hammond, *et al.*, “Unifying and accelerating level-set and density-based topology optimization by subpixel-smoothed projection,” *Opt. Express* **33**, 33620 (2025).
  53. A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, “Automatic differentiation in machine learning: A survey,” *J. Mach. Learn. Res.* **18**, 1–43 (2018).
  54. A. Griewank and A. Walther, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation* (SIAM, Philadelphia, 2008), 2nd ed.
  55. M. Innes, “Don’t unroll adjoint: differentiating SSA-form programs,” [arXiv:1810.07951](https://arxiv.org/abs/1810.07951) (2018).
  56. K. Svanberg, “A class of globally convergent optimization methods based on conservative convex separable approximations,” *SIAM J. on Optim.* **12**, 19 (2002).
  57. S. G. Johnson, “The NLopt nonlinear-optimization package,” [github.com/stevengj/nlopt](https://github.com/stevengj/nlopt) (2007).
  58. A. Mutapcic, S. Boyd, A. Farjadpour, *et al.*, “Robust design of slow-light tapers in periodic waveguides,” *Eng. Optim.* **41**, 365–384 (2009).
  59. A. Blake and A. Zisserman, *Visual Reconstruction* (MIT Press, 1987).
  60. X. Lin, Z. Yang, X. Zhang, and Q. Zhang, “Continuation path learning for homotopy optimization,” in *Proc. 40th Intl. Conf. Machine Learning*, vol. 202 of *Proceedings of Machine Learning Research* (PMLR, 2023), pp. 21288–21311.
  61. T. J. Authors, “The autodiff cookbook,” [https://docs.jax.dev/en/latest/notebooks/autodiff\\_cookbook.html](https://docs.jax.dev/en/latest/notebooks/autodiff_cookbook.html) (2025).
  62. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, 2007), 3rd ed.
  63. S. He, Y. Shi, E. Jonsson, and J. R. Martins, “Eigenvalue problem derivatives computation for a complex matrix using the adjoint method,” *Mech. Syst. Signal Process.* **185**, 109717 (2023).
  64. B. Li and G. J. Kennedy, “Adjoint methods for computing derivatives of functions of eigenvectors using shift-and-invert preconditioning,” *Struct. Multidiscip. Optim.* **68** (2024).
  65. R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, 2012).
  66. A. Montoisson and D. Orban, “Krylov.jl: A Julia basket of hand-picked Krylov methods,” *J. Open Source Softw.* **8**, 5187 (2023).
  67. S. G. Johnson, “Notes on perfectly matched layers (PMLs),” [arXiv:2108.05348](https://arxiv.org/abs/2108.05348) (2021).
  68. W. C. Chew and W. H. Weedon, “A 3d perfectly matched medium from modified Maxwell’s equations with stretched coordinates,” *Microw. Opt. Technol. Lett.* **7**, 599–604 (1994).
  69. A. Oskooi and S. G. Johnson, “Distinguishing correct from incorrect PML proposals and a corrected unsplit PML for anisotropic, dispersive media,” *J. Comput. Phys.* **230**, 2369–2377 (2011).