

# Formulation for scalable optimization of microcavities via the frequency-averaged local density of states

Xiangdong Liang\* and Steven G. Johnson

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

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

**Abstract:** We present a technique for large-scale optimization of optical microcavities based on the frequency-averaged local density of states (LDOS), which circumvents computational difficulties posed by previous eigenproblem-based formulations and allows us to perform full topology optimization of three-dimensional (3d) leaky cavity modes. We present theoretical results for both 2d and fully 3d computations in which *every pixel* of the design pattern is a degree of freedom (“topology optimization”), e.g. for lithographic patterning of dielectric slabs in 3d. More importantly, we argue that such optimization techniques can be applied to design cavities for which (unlike silicon-slab single-mode cavities) hand designs are difficult or unavailable, and in particular we design minimal-volume *multi-mode* cavities (e.g. for nonlinear frequency-conversion applications).

© 2013 Optical Society of America

**OCIS codes:** (120.4570) Optical design of instruments; (140.3945) Microcavities; (230.5750) Resonators.

---

## References and links

1. M. Nomura, K. Tanabe, S. Iwamoto, and Y. Arakawa, “High-Q design of semiconductor-based ultrasmall photonic crystal nanocavity,” *Opt. Express* **18**, 8144–8150 (2010).
2. B. Bourdin, “Filters in topology optimization,” *Int. J. Numer. Methods Eng.* **50**, 2143–2158 (2001).
3. M. P. Bendsoe and O. Sigmund, *Topology Optimization: Theory, Methods and Applications*, 2nd ed. (Springer, 2003).
4. M. Y. Wang, X. Wang, and D. Guo, “A level set method for structural topology optimization,” *Comput. Methods Appl. Mech. Eng.* **192**, 227–246 (2003).
5. G. Allaire, F. Jouve, and A.-M. Toader, “Structural optimization using sensitivity analysis and a level-set method,” *J. Comput. Phys.* **194**, 363–393 (2004).
6. O. Sigmund, “Manufacturing tolerant topology optimization,” *Acta Mech. Sinica* **25**, 227–239 (2009).
7. E. Andreassen, A. Clausen, M. Schevenels, B. Lazarov, and O. Sigmund, “Efficient topology optimization in MATLAB using 88 lines of code,” *Struct. Multidiscip. Optim.* **43**, 1–16 (2011).
8. J. S. Jensen and O. Sigmund, “Topology optimization for nano-photonics,” *Laser Photonics Rev.* **5**, 308–321 (2011).
9. J. Bravo-Abad, A. Rodriguez, P. Bermel, S. G. Johnson, J. D. Joannopoulos, and M. Soljacic, “Enhanced nonlinear optics in photonic-crystal microcavities,” *Opt. Express* **15**, 16161–16176 (2007).
10. R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Academic, 2008).
11. H. Hashemi, C. W. Qiu, A. P. McCauley, J. D. Joannopoulos, and S. G. Johnson, “A diameter–bandwidth product limitation of isolated-object cloaking,” *Phys. Rev. A* **86**, 013804 (2012).
12. D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. (Athena Scientific, 1999).
13. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics (Wiley Series in Pure and Applied Optics)*, 2nd ed. (Wiley-Interscience, 2007).
14. J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University, 2008).

15. D. G. Rabus, *Integrated Ring Resonators*, Vol. 127 of *Springer Series in Optical Sciences* (Springer, 2007).
16. S. Lin, E. Schonbrun, and K. Crozier, "Optical manipulation with planar silicon microring resonators," *Nano Lett.* **10**, 2408–2411 (2010).
17. J. Vuckovic, M. Loncar, H. Mabuchi, and A. Scherer, "Optimization of the Q factor in photonic crystal microcavities," *IEEE J. Quantum Electron.* **38**, 850–856 (2002).
18. C. W. Wong, P. T. Rakich, S. G. Johnson, M. Qi, H. I. Smith, E. P. Ippen, L. C. Kimerling, Y. Jeon, G. Barbastathis, and S. G. Kim, "Strain-tunable silicon photonic band gap microcavities in optical waveguides," *Appl. Phys. Lett.* **84**, 1242–1244 (2004).
19. Y. Akahane, T. Asano, B.-S. Song, and S. Noda, "High-Q photonic nanocavity in a two-dimensional photonic crystal," *Nature* **425**, 944–947 (2003).
20. Y. Akahane, T. Asano, B.-S. Song, and S. Noda, "Fine-tuned high-Q photonic-crystal nanocavity," *Opt. Express* **13**, 1202–1214 (2005).
21. B.-S. Song, S. Noda, T. Asano, and Y. Akahane, "Ultra-high-Q photonic double-heterostructure nanocavity," *Nat. Mater.* **4**, 207–210 (2005).
22. D. C. Dobson and F. Santosa, "Optimal Localization of Eigenfunctions in an Inhomogeneous Medium," *SIAM J. Appl. Math.* **64**, 762–774 (2004).
23. C.-Y. Kao and F. Santosa, "Maximization of the quality factor of an optical resonator," *Wave Motion* **45**, 412–427 (2008).
24. W. R. Frei, H. T. Johnson, and K. D. Choquette, "Optimization of a single defect photonic crystal laser cavity," *J. Appl. Phys.* **103**, 033102 (2008).
25. J. Lu and J. Vuckovic, "Inverse design of nanophotonic structures using complementary convex optimization," *Opt. Express* **18**, 3793–3804 (2010).
26. A. W. Snyder and J. Love, *Optical Waveguide Theory, Science Paperbacks, 190* (Springer, 1983).
27. R. Coccioli, M. Boroditsky, K. W. Kim, Y. Rahmat-Samii, and E. Yablonovitch, "Smallest possible electromagnetic mode volume in a dielectric cavity," *IEE Proceedings - Optoelectronics* **145**, 391–397 (1998).
28. E. M. Purcell, "Spontaneous emission probabilities at radio frequencies," *Phys. Rev.* **69**, 674 (1946).
29. L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University, 2006).
30. A. F. Koenderink, "On the use of Purcell factors for plasmon antennas," *Opt. Lett.* **35**, 4208–4210 (2010).
31. B. Zhen, S.-L. Chua, J. Lee, A. W. Rodriguez, X. Liang, S. G. Johnson, J. D. Joannopoulos, M. Soljačić, and O. Shapira, "Enabling enhanced emission and low-threshold lasing of organic molecules using special Fano resonances of macroscopic photonic crystals," *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13711–13716 (2013).
32. E. A. J. Marcatili, "Bends in optical dielectric guides," *Bell Syst. Tech. J.* **48**, 2103–2132 (1969).
33. F. Ladouceur, "Roughness, inhomogeneity, and integrated optics," *J. Lightwave Technol.* **15**, 1020–1025 (1997).
34. V. S. Ilchenko, P. S. Volikov, V. L. Velichansky, F. Treussart, V. Lefèvre-Seguin, J. M. Raimond, and S. Haroche, "Strain-tunable high-Q optical microsphere resonator," *Opt. Commun.* **145**, 86–90 (1998).
35. M. Soltani, S. Yegnanarayanan, and A. Adibi, "Ultra-high Q planar silicon microdisk resonators for chip-scale silicon photonics," *Opt. Express* **15**, 4694–4704 (2007).
36. L. N. Trefethen and D. Bau, *Numerical Linear Algebra* (SIAM, 1997).
37. K. Inoue and K. Ohtaka, *Photonic Crystals: Physics, Fabrication and Applications, Springer Series in Optical Sciences* (Springer, 2010).
38. A. Oskooi and S. G. Johnson, "Electromagnetic wave source conditions," in *Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology*, A. Taflov, A. Oskooi, and S. G. Johnson, eds. (Artech, 2013), Chap. 4, pp. 65–100.
39. J.-M. Gerard and B. Gayral, "Strong Purcell effect for InAs quantum boxes in three-dimensional solid-state microcavities," *J. Lightwave Technol.* **17**, 2089–2095 (1999).
40. O. J. F. Martin and N. B. Piller, "Electromagnetic scattering in polarizable backgrounds," *Phys. Rev. E* **58**, 3909–3915 (1998).
41. G. D'Aguzzo, N. Mattiucci, M. Centini, M. Scalora, and M. J. Bloemer, "Electromagnetic density of modes for a finite-size three-dimensional structure," *Phys. Rev. E* **69**, 057601 (2004).
42. J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (John Wiley, 1975).
43. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge University, 2007).
44. M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory (Frontiers in Physics)* (Westview, 1995).
45. S. G. Johnson, "Numerical methods for computing Casimir interactions," in *Casimir Physics*, Vol. 834 of *Lecture Notes in Physics*, D. Dalvit, P. Milonni, D. Roberts, and F. da Rosa, eds. (Springer, 2011), Chap. 6.
46. L. D. Landau, L. P. Pitaevskii, and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed. (Butterworth-Heinemann, 1984).
47. W. C. Chew, *Waves and Fields in Inhomogeneous Media* (IEEE, 1995).
48. N. A. P. Nicorovici, R. C. McPhedran, and L. C. Botten, "Relative local density of states for homogeneous lossy materials," *Physica B* **405**, 2915–2919 (2010).
49. L. V. Ahlfors, *Complex Analysis*, 3rd ed. (McGraw-Hill, 1978).

50. A. W. Rodriguez, A. P. McCauley, J. D. Joannopoulos, and S. G. Johnson, "Theoretical ingredients of a Casimir analog computer," *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9531–9536 (2010).
51. X. Liang, Ph.D. thesis, Massachusetts Institute of Technology, 2013.
52. S. Scheel, L. Knöll, and D. G. Welsch, "Spontaneous decay of an excited atom in an absorbing dielectric," *Phys. Rev. A* **60**, 4094–4104 (1999).
53. C. Van Vlack and S. Hughes, "Finite-difference time-domain technique as an efficient tool for calculating the regularized Green function: applications to the local-field problem in quantum optics for inhomogeneous lossy materials," *Opt. Lett.* **37**, 2880–2882 (2012).
54. G. Strang, *Computational Science and Engineering* (Wellesley-Cambridge, 2007).
55. A. Christ and H. L. Hartnagel, "Three-dimensional finite-difference method for the analysis of microwave-device embedding," *IEEE Trans. Microwave Theory* **35**, 688–696 (1987).
56. K. Yasumoto, *Electromagnetic Theory and Applications for Photonic Crystals, Optical Science and Engineering* (CRC, 2005).
57. 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).
58. T. A. Davis, *Direct Methods for Sparse Linear Systems (Fundamentals of Algorithms)* (SIAM, 2006).
59. S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith, "Efficient Management of Parallelism in Object Oriented Numerical Software Libraries," In *Modern Software Tools in Scientific Computing*, E. Arge, A. M. Bruaset, and H. P. Langtangen, eds. (Birkhäuser, 1997).
60. S. Balay, J. Brown, K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang, "PETSc Users Manual," Technical Report No. ANL-95/11 - Revision 3.3, Argonne National Laboratory (2012).
61. S. Balay, J. Brown, K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang, "PETSc Web page," <http://www.mcs.anl.gov/petsc>, 2012.
62. P. Hénon, P. Ramet, and J. Roman, "PaStiX: a high-performance parallel direct solver for sparse symmetric positive definite systems," *Parallel Computing* **28**, 301–321 (2002).
63. S. G. Johnson, "The NLOpt nonlinear-optimization package," <http://ab-initio.mit.edu/nlopt>.
64. D. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Program.* **45**, 503–528 (1989).
65. K. Svanberg, "A class of globally convergent optimization methods based on conservative convex separable approximations," *SIAM J. Optimiz.* **12**, 555–573 (2002).
66. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, 2000).
67. A. Mutapcic, S. Boyd, A. Farjadpour, S. G. Johnson, and Y. Avniel, "Robust design of slow-light tapers in periodic waveguides," *Eng. Optimiz.* **41**, 365–384 (2009).
68. A. Oskooi, A. Mutapcic, S. Noda, J. D. Joannopoulos, S. P. Boyd, and S. G. Johnson, "Robust optimization of adiabatic tapers for coupling to slow-light photonic-crystal waveguides," *Opt. Express* **20**, 21558–21575 (2012).
69. Y. Xu, W. Liang, A. Yariv, J. G. Fleming, and S.-Y. Lin, "High-quality-factor Bragg onion resonators with omnidirectional reflector cladding," *Opt. Lett.* **28**, 2144–2146 (2003).
70. Z. Artstein, "Discrete and continuous bang-bang and facial spaces or: Look for the extreme points," *SIAM Rev.* **22**, 172–185 (1980).
71. A. F. Oskooi, Ph.D. thesis, Massachusetts Institute of Technology, 2010.
72. J. Lu, S. Boyd, and J. Vuckovic, "Inverse design of a three-dimensional nanophotonic resonator," *Opt. Express* **19**, 10563–10570 (2011).
73. B. Osting and M. I. Weinstein, "Long-lived scattering resonances and Bragg structures," *SIAM J. Appl. Math.* **73**, 827–852 (2013).
74. T. Inui, Y. Tanabe, and Y. Onodera, *Group Theory and Its Applications in Physics* (Springer, 1996).
75. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University, 2004).
76. P. D. Drummond, K. J. McNeil, and D. F. Walls, "Non-equilibrium transitions in sub/second harmonic generation," *Opt. Acta* **27**, 321–335 (1980).
77. L.-A. Wu, M. Xiao, and H. J. Kimble, "Squeezed states of light from an optical parametric oscillator," *J. Opt. Soc. Am. B* **4**, 1465–1475 (1987).
78. Z. Y. Ou and H. J. Kimble, "Enhanced conversion efficiency for harmonic generation with double resonance," *Opt. Lett.* **18**, 1053–1055 (1993).
79. A. Rodriguez, M. Soljacic, J. D. Joannopoulos, and S. G. Johnson, " $\chi^{(2)}$  and  $\chi^{(3)}$  harmonic generation at a critical power in inhomogeneous doubly resonant cavities," *Opt. Express* **15**, 7303–7318 (2007).
80. I. B. Burgess, Y. Zhang, M. W. McCutcheon, A. W. Rodriguez, J. Bravo-Abad, S. G. Johnson, and M. Loncar, "Design of an efficient terahertz source using triply resonant nonlinear photonic crystal cavities," *Opt. Express* **17**, 20099–20108 (2009).
81. Z.-F. Bi, A. W. Rodriguez, H. Hashemi, D. Duchesne, M. Loncar, K.-M. Wang, and S. G. Johnson, "High-efficiency second-harmonic generation in doubly-resonant  $\chi^{(2)}$  microring resonators," *Opt. Express* **20**, 7526–7543 (2012).
82. M. W. McCutcheon and M. Loncar, "Design of a silicon nitride photonic crystal nanocavity with a Quality factor

- of one million for coupling to a diamond nanocrystal,” *Opt. Express* **16**, 19136–19145 (2008).
83. A. R. Md Zain, N. P. Johnson, M. Sorel, and R. M. De La Rue, “Ultra high quality factor one dimensional photonic crystal/photonic wire micro-cavities in silicon-on-insulator (SOI),” *Opt. Express* **16**, 12084–12089 (2008).
  84. Z. M. Meng, F. Qin, Y. Liu, and Z. Y. Li, “High-Q microcavities in low-index one-dimensional photonic crystal slabs based on modal gap confinement,” *J. Appl. Phys.* **109**, 043107 (2011).
  85. J. T. Robinson, C. Manolatu, L. Chen, and M. Lipson, “Ultrasmall mode volumes in dielectric optical micro-cavities,” *Phys. Rev. Lett.* **95**, 143901 (2005).
  86. M. Nomura, “GaAs-based air-slot photonic crystal nanocavity for optomechanical oscillators,” *Opt. Express* **20**, 5204–5212 (2012).
  87. S. Kita, K. Nozaki, S. Hachuda, H. Watanabe, Y. Saito, S. Otsuka, T. Nakada, Y. Arita, and T. Baba, “Photonic Crystal Point-Shift Nanolasers With and Without Nanoslots Design, Fabrication, Lasing, and Sensing Characteristics,” *IEEE J. Sel. Top. Quantum Electron.* **17**, 1632–1647 (2011).
  88. Y. Li, J. Zheng, J. Gao, J. Shu, M. S. Aras, and C. W. Wong, “Design of dispersive optomechanical coupling and cooling in ultrahigh-Q/V slot-type photonic crystal cavities,” *Opt. Express* **18**, 23844–23856 (2010).
  89. O. Sigmund and J. Petersson, “Numerical instabilities in topology optimization: A survey on procedures dealing with checkerboards, mesh-dependencies and local minima,” *Struct. Multidiscip. Optim.* **16**, 68–75 (1998).
  90. S. J. Osher and R. P. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces* (Springer, 2002).
  91. 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**, 235–244 (2005).
  92. J. K. Guest, J. H. Prévost, and T. Belytschko, “Achieving minimum length scale in topology optimization using nodal design variables and projection functions,” *Int. J. Numer. Methods Eng.* **61**, 238–254 (2004).
  93. O. Sigmund, “Morphology-based black and white filters for topology optimization,” *Struct. Multidiscip. Optim.* **33**, 401–424 (2007).
  94. S. Xu, Y. Cai, and G. Cheng, “Volume preserving nonlinear density filter based on heaviside functions,” *Struct. Multidiscip. Optim.* **41**, 495–505 (2010).
  95. F. Wang, B. Lazarov, and O. Sigmund, “On projection methods, convergence and robust formulations in topology optimization,” *Struct. Multidiscip. Optim.* **43**, 767–784 (2011).
  96. A. Ben-Tal, L. El Ghaoui, and A. S. Nemirovski, *Robust Optimization* (Princeton University, 2009).
  97. D. Bertsimas, O. Nohadani, and K. M. Teo, “Robust optimization for unconstrained simulation-based problems,” *Oper. Res.* **58**, 161–178 (2010).
  98. F. Wang, J. S. Jensen, and O. Sigmund, “Robust topology optimization of photonic crystal waveguides with tailored dispersion properties,” *J. Opt. Soc. Am. B* **28**, 387–397 (2011).
  99. H. Men, R. M. Freund, N. C. Nguyen, J. Saa-Seoane, and J. Peraire, “Fabrication-adaptive optimization with an application to photonic crystal design,” arXiv:1307.5571.
  100. A. Faraon, C. Santori, Z. Huang, V. M. Acosta, and R. G. Beausoleil, “Coupling of nitrogen-vacancy centers to photonic crystal cavities in monocrystalline diamond,” *Phys. Rev. Lett.* **109**, 033604 (2012).
  101. L. Li, M. Trusheim, O. Gaathon, K. Kisslinger, C.-J. Cheng, M. Lu, D. Su, X. Yao, H.-C. Huang, I. Bayn, A. Wolcott, R. M. Osgood Jr., and D. Englund, “Reactive ion etching: Optimized diamond membrane fabrication for transmission electron microscopy,” *J. Vac. Sci. Technol. B* **31**, 06FF01 (2013).
  102. Z. Yu, A. Raman, and S. Fan, “Fundamental limit of light trapping in grating structures,” *Opt. Express* **18**, A366–A380 (2010).
  103. Z. Yu, A. Raman, and S. Fan, “Fundamental limit of nanophotonic light trapping in solar cells,” *Proc. Natl. Acad. Sci. U. S. A.* **107**, 17491–17496 (2010).

---

## 1. Introduction

In this paper, we present a new technique for large-scale optimization of optical microcavities based on the frequency-averaged local density of states (LDOS), which circumvents computational difficulties posed by previous eigenproblem-based formulations and allows us to perform full topology optimization of three-dimensional (3d) leaky cavity modes. Essentially, this technique allows us to minimize modal volume  $V$  for a given cavity quality factor  $Q$ , or equivalently a given bandwidth, by solving a single *complex-frequency* scattering problem at each optimization step, which is both computationally easier than solving an eigenproblem and avoids the difficulty of selecting which eigenvalue to optimize. We present proof-of-concept results in both 2d and 3d computations in which *every pixel* of the design pattern is a degree of freedom (“topology optimization”), e.g. for lithographic patterning of dielectric slabs in 3d. For a design  $Q$  of  $\approx 10^4$  in 3d for silicon slabs (index  $n = 3.52$ ), we obtain a modal volume  $V = 0.06(\lambda/n)^3$  (relative to the vacuum wavelength  $\lambda$ ), which is  $4\times$  smaller than the best

comparable- $Q$  volume found in the literature [1]. Here, the focus is on illustrating the technique rather than on designing practical cavities (for which many good designs are already available in the single-mode silicon case), so we do not incorporate regularizations [2–8] to force the optimization to find easily fabricated structures. However, our results are still useful in establishing theoretical bounds, showing that significant room for improvement remains even for silicon cavities. More importantly, we show that such optimization techniques can be applied to design cavities for which (unlike silicon-slab single-mode cavities) hand designs are difficult or unavailable, and in particular we design proof-of-concept minimal-volume *multi-mode* cavities (e.g. for nonlinear frequency-conversion applications [9, 10]) in Secs. 8.2.3 and 8.3. We find that the optimum doubly degenerate cavity appears to be three-fold symmetric, while two-frequency cavities have more complex shapes depending on the polarization, although further work is needed to obtain manufacturable structures. Theoretical optimization is also useful to investigate the tradeoff between the size of the design region and the maximum attainable  $Q$  in slab-like situations (Sec. 8.4), and find it to be roughly exponential. All of these results are enabled by a sequence of mathematical transformations of the original cavity-design problem, depicted in Fig. 1: from the Purcell factor  $Q/V$  to the more well-posed problem of minimizing  $V$  for a given  $Q$  (Sec. 2), from the eigenproblem to a scattering problem via the LDOS (Sec. 3), from a frequency-averaged LDOS to a single *complex*-frequency scattering problem [11] via contour integration (Secs. 4–5), and finally from maximizing LDOS to minimizing  $1/\text{LDOS}$  in order to circumvent optimization problems that arise for sharply peaked objectives (Sec. 7). Even though the optimization problem is nonlinear and probably non-convex [12], we obtain similar optima for many initial structures (including vacuum or random pixels), suggesting that the result may be close to a global optimum.

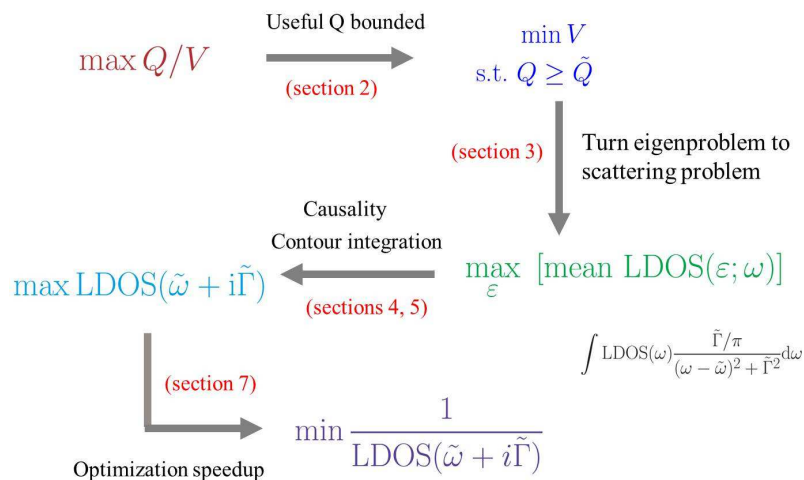


Fig. 1. Starting with the naive objective of maximizing a microcavity’s Purcell factor  $Q/V$ , we perform a sequence of transformations of the problem in order to make it well posed and tractable. Here, we give a schematic diagram of each transformation, along with the corresponding section of the paper in which they are discussed.

Microcavity design, which seeks to confine a cavity “mode” for a long time (dimensionless lifetime  $Q$ ) in a small “modal volume”  $V$ , has a long history [13, 14], and until recently has been dominated by hand designs, from ring resonators [15, 16] to photonic-crystal slab cavities [17, 18], in which the parameters are manually tweaked to obtain the desired results (or in some cases computer optimization is performed over a handful of parameters [19–21]). These



manual designs have been tremendously successful, attaining experimentally verified  $Q$  of  $10^6$  or more in modal volumes only a few wavelengths in diameter [19–21]. However, fundamental questions remain: how close are the existing designs to the theoretical optima (e.g. the minimal  $V$  for any given  $Q$ ), and do the optimal designs resemble the hand designs or are they entirely different? (As discussed in Sec. 2, the question is *not* what is the maximum  $Q$  or  $Q/V$ , because those questions have a trivial theoretical answer:  $\infty$ .) Furthermore, while design of single-mode cavities in silicon slabs has been heavily studied and many design heuristics are known, very different designs and grueling effort may be required in radically different circumstances, from new materials to new design goals such as multi-mode cavities for nonlinear optics [10]. Large-scale “topology” optimization, in which the design pattern is completely determined by computational search (often with thousands of free parameters) with little or no *a priori* information, offers a different route to addressing these questions. Although several authors have developed topology-optimization techniques for microcavity design [22–25], most of that previous work was limited to 2d calculations [22, 23] and none of the previous work fully addressed the design tradeoff between  $Q$  and  $V$ . Some topology optimization dealt with lossless systems and avoided  $Q$  entirely [22], or maximized a 2d  $Q$  without controlling  $V$  [23] (which we argue in Sec. 2 leads to an ill-posed optimization problem). Other work used a “2.5d” heuristic for the cavity radiation loss [25], which permitted 2d calculations and led to intriguing designs, but limits the attainable  $Q$  (to  $< 10^4$ ) because there appears to be no systematic way to improve this heuristic loss estimate. One group did perform fully 3d calculations with absorbing boundaries to capture radiation loss [24], but limited their degrees of freedom to a small region inside a photonic crystal and included an unphysical absorbing material within the cavity itself, leading to a heuristic objective whose calculation does not appear to have a rigorous quantitative relationship to key cavity properties. As we review in Sec. 3, maximizing LDOS corresponds mathematically to maximizing the power radiated by a dipole current source inside the cavity. Naively, this may seem backwards: isn’t the purpose of a cavity to *minimize* radiated power? However, that intuition stems from the situation in which the energy density *inside* the cavity is held constant: there is an equation  $Q = \omega U/P$  relating the cavity  $Q$  and frequency  $\omega$  with the radiated power  $P$  and the confined energy  $U$  [14], so that  $P \sim 1/Q$  for fixed  $U$ . In our case, however, we are holding the *current amplitude* in the cavity fixed, in which case the radiated power  $P \sim Q$  (as reviewed in Sec. 3) while one can show from a coupled-mode framework [14] that  $U \sim Q^2$ . Conversely, minimizing radiated power for a fixed current would correspond to *minimizing* LDOS, which would result in the *absence* of resonances. (Interestingly, Frei et al. [24] actually *minimized* radiated power from a fixed-current dipole, but their introduction of a heuristic absorbing region inside the cavity apparently compensated for this inverted objective and resulted in a resonant mode, albeit a mode optimizing an unclear objective.) On the other hand, our work adapts two crucial ideas from [24]: solving a scattering problem rather than an eigenproblem, and optimizing over a finite bandwidth. Nor did any previous work, to our knowledge, address topology optimization of multi-mode cavities.

## 2. Eigenproblem formulation

There are two key figures of merit for a resonant mode  $\mathbf{E}^n(\mathbf{x})$  of a cavity: quality factor  $Q$  and modal volume  $V$ . The quality factor  $Q$  is a dimensionless lifetime, and  $1/Q$  is a dimensionless decay rate [14]. Mathematically,  $Q$  is related to the frequency-domain Maxwell eigenvalue problem:

$$\nabla \times \frac{1}{\mu(\mathbf{x})} \nabla \times \mathbf{E}^n(\mathbf{x}) = \omega_n^2 \epsilon(\mathbf{x}) \mathbf{E}^n(\mathbf{x}) \quad (1)$$

with radiation boundary conditions. Because of the lossy boundary conditions, the eigenproblem is non-Hermitian and the eigenvalues are complex. The  $Q$  for the mode  $\mathbf{E}^n(\mathbf{x})$  with eigen-

frequency  $\omega_n$  [14] is

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

Technically, if one considers an infinite open system (as opposed to a finite system with some absorbing boundary layer), a number of mathematical subtleties arise because the leaky “modes” are not true normalizable eigenfunctions, but are rather related to the residues of poles in the Green’s function [26]; however, we will circumvent all of these difficulties in this paper by transition to a Green’s-function LDOS approach in Sec. 3 that does not deal explicitly with eigenvalue problems. The modal volume  $V$  [27], defined as

$$V = \frac{\int \varepsilon(\mathbf{x}) |\mathbf{E}^n(\mathbf{x})|^2 d\mathbf{x}}{\max\{\varepsilon(\mathbf{x}) |\mathbf{E}^n(\mathbf{x})|^2\}}, \quad (3)$$

is a measure of the volume within which the mode is confined. (Although Eq. (3) is a standard expression, it should really be viewed as the modal volume in the limit of a lossless cavity. For a lossy cavity,  $\int \varepsilon(\mathbf{x}) |\mathbf{E}^n(\mathbf{x})|^2 d\mathbf{x}$  does not converge in an infinite open system; a more rigorous treatment is to define the LDOS in terms of the Green’s function as in Sec. 3 and obtain the Purcell factor from the ratio with the LDOS of the homogeneous medium.) The Purcell factor [28], which describes the enhancement of spontaneous emission rate, can be written as [29, 30]

$$\frac{3Q}{4\pi^2 V} \left(\frac{\lambda}{n}\right)^3. \quad (4)$$

Here  $\lambda$  is the vacuum wavelength and  $n$  is the index of refraction.

For applications with light-matter interactions (such as lasers, sensors, and nonlinear frequency converters), maximal lifetime  $Q$  and minimal modal volume  $V$  are desirable [31]. It is therefore tempting to use the Purcell factor in Eq. (4) or  $Q/V$  as the figure of merit for cavity optimization. Unfortunately, maximizing  $Q/V$  leads to an ill-posed problem, because the maximum of  $Q/V$  is  $\infty$ ; for example,  $Q/V$  grows exponentially with radius for a ring resonator [32]. In practice, any optimization in a finite computation cell will obtain a finite  $Q$  and  $V$  [23], but the values are then just an artifact of the finite computational domain; in this sense, maximizing  $Q/V$  is not well-posed because the solution does not converge as one increases the size of the computational domain.

In practice, however, there is an upper bound on the useful  $Q$  for two reasons. First, besides the intrinsic radiation loss ( $Q_{\text{rad}}$ ) in a cavity, there are also radiation losses due to surface roughness ( $Q_{\text{roughness}}$ ) and material absorption ( $Q_{\text{absorption}}$ ). The total loss rate  $1/Q_{\text{total}}$  is the sum of these three effects [14]:

$$\frac{1}{Q_{\text{total}}} = \frac{1}{Q_{\text{rad}}} + \frac{1}{Q_{\text{roughness}}} + \frac{1}{Q_{\text{absorption}}}. \quad (5)$$

In real applications, therefore, the  $Q_{\text{total}}$  cannot be arbitrarily large. For example, in integrated optics it is difficult to get  $Q_{\text{total}}$  more than a few million due to surface roughness [33], although microdisk and microsphere resonators can achieve higher  $Q$  by delocalizing the mode away from the surface [34, 35]. Second, there is another quality factor in the system. For any cavity-based device, the cavity is always intentionally coupled to some channels (e.g., waveguides) to get light in and out. That coupling process will be described by its own lifetime  $Q_{\text{coupling}}$ . It turns out that the losses in such a coupled device are proportional to  $Q_{\text{coupling}}/Q_{\text{total}}$  [14]. Once these losses are decreased below the desired loss budget, it often does not matter in practice if one decreases them further.

A better optimization problem might be, instead, to maximize  $Q/V$  subject to  $Q = \tilde{Q}$ , or alternatively:

$$\begin{aligned} \min V \\ \text{s.t. } Q \geq \tilde{Q}, \end{aligned} \quad (6)$$

where  $\tilde{Q}$  is determined by the bandwidth and loss tolerance of applications. By solving the non-Hermitian Maxwell eigenproblem Eq. (1), one could obtain  $Q$  and  $V$  from eigenvectors  $\mathbf{E}^n(\mathbf{x})$  and eigenvalues  $\omega_n$  through Eq. (2) and Eq. (3). Then a natural question to ask is which eigenvalue one should optimize. In practice, one has some design frequency  $\tilde{\omega}$  given by the application, so one could optimize the eigenvalue closest to  $\tilde{\omega}$ . (Equivalently, thanks to the scale-invariance of Maxwell's equations [14], we choose units so that  $\tilde{\omega} = 2\pi$ , in which case the main computational choice is the resolution, i.e. the number of pixels per wavelength  $\lambda$ , and in some cases a slab thickness/ $\lambda$ .) However, asking for the mode closest to  $\tilde{\omega}$  leads to discontinuities: as the structure changes during optimization, the mode closest to  $\tilde{\omega}$  will tend to hop discontinuously. Although there are some ways to deal with this [22], the problem becomes worse when one simulates the radiation loss in a patterned dielectric slab, because in this case the finite cell is approximating a continuum of radiation modes above and below the slab. As a result, there are more and more closely spaced modes as the cell size increases. Hence, we want to circumvent this difficulty by adopting a new approach: turning the eigenproblem into a linear scattering problem. This also reduces the computational expense, because we will now require only a single linear solve per structure, whereas finding eigenvalues in the interior of a spectrum (e.g. by the shift-and-invert algorithm) requires many linear solves [36].

### 3. LDOS formulation

The well-known Purcell factor ( $Q/V$  and variations thereof) is, in fact, only an *approximation* of a more fundamental quantity, the local density of states (LDOS), which is defined in terms of the Green's function of the system (and can be related to a "density of modes" per unit frequency per unit volume) [37]. Paradoxically, the *exact* LDOS is *easier* to compute than  $Q/V$ , because the former involves only a scattering problem (a linear system of equations) whereas the latter involves a non-Hermitian eigenproblem (for a leaky mode in the interior of the spectrum). In this section, we briefly review the definition of the LDOS and how it relates to  $Q/V$ , and in the next section we describe how the LDOS can be used as the objective for a well-posed cavity optimization problem.

In particular, we begin with the per-polarization LDOS (partial LDOS), denoted by  $\text{LDOS}_j(\omega, \mathbf{x}')$  for a polarization in direction  $j$ , which is proportional to the power radiated by a dipole in direction (unit vector)  $\hat{\mathbf{e}}_j$  at position  $\mathbf{x}'$  with a frequency  $\omega$ , i.e. a current  $\mathbf{J} \sim \hat{\mathbf{e}}_j e^{-i\omega t} \delta(\mathbf{x} - \mathbf{x}')$ . (The total LDOS is simply  $\sum_j \text{LDOS}_j$ , proportional to power radiated by a randomly oriented dipole.) This is a key physical quantity because quantum fluctuations, e.g. spontaneous emission or thermal fluctuations, can be viewed semiclassically as dipole currents, and the LDOS yields the enhancement of these fluctuations by any given geometry. For example, the spontaneous emission rate is proportional to the LDOS [29, 38], and therefore the ratio of LDOS between two systems indicates enhancement or suppression of spontaneous emission. The Purcell factor  $Q/V$  is an approximation to this LDOS enhancement for a high- $Q$  microcavity, assuming that the emitter (e.g. atom or quantum dot) is positioned and oriented at the location of peak  $\text{LDOS}_j$  in the cavity [29, 30, 38, 39].

These quantities can be defined more precisely as follows [29, 37, 38, 40, 41]. Poynting's theorem [42] implies that the power radiated by a dipole  $\mathbf{J}(\mathbf{x})$  is

$$P_j(\omega, \mathbf{x}') = -\frac{1}{2} \int \text{Re} [\mathbf{J}^*(\mathbf{x}) \cdot \mathbf{E}(\mathbf{x})] d\mathbf{x}, \quad (7)$$



where  $\mathbf{E}(\mathbf{x})$  is the total electric field solving the frequency-domain scattering problem

$$\begin{aligned}\mathcal{M}(\boldsymbol{\varepsilon}, \omega)\mathbf{E}(\mathbf{x}) &= i\omega\mathbf{J}(\mathbf{x}) \\ \mathcal{M}(\boldsymbol{\varepsilon}, \omega) &= \nabla \times \frac{1}{\mu(\mathbf{x})} \nabla \times - \boldsymbol{\varepsilon}(\mathbf{x})\omega^2 \\ \mathbf{J}(\mathbf{x}) &= \delta(\mathbf{x} - \mathbf{x}')\hat{\mathbf{e}}_j.\end{aligned}\quad (8)$$

This yields the following well known [29, 37, 38, 40, 41] definition of LDOS<sub>*j*</sub>:

$$\text{LDOS}_j(\omega, \mathbf{x}') = \frac{12}{\pi} P_j(\omega, \mathbf{x}') = -\frac{6}{\pi} \text{Re} \left[ \int \mathbf{J}^*(\mathbf{x}) \cdot \mathbf{E}(\mathbf{x}) d\mathbf{x} \right], \quad (9)$$

where the  $12/\pi$  factor is a conventional normalization that arises from the dual interpretation of the LDOS as a local “density” of eigenstates [29, 38]. (The normalization is irrelevant for optimization or for evaluating LDOS ratios in different systems to obtain enhancement factors.)

In the limit of a low-loss cavity, in which the LDOS is dominated by the contribution of a single pole in Maxwell’s equations (a single “resonant mode”), one can derive the approximation [38]

$$\text{LDOS}_j(\omega, \mathbf{x}') = \frac{6}{\pi\omega\varepsilon(\mathbf{x}')} \frac{Q}{V}, \quad (10)$$

and by taking the ratio of this quantity with the LDOS of a homogeneous medium with index  $n = \sqrt{\varepsilon}$  one [29] obtains the traditional Purcell factor Eq. (4). In particular, the LDOS in a microcavity resonating a frequency  $\tilde{\omega}(1 + i/2\tilde{Q})$  is approximately in the form of a Lorentzian peak centered at  $\tilde{\omega}$  with a bandwidth  $\tilde{\omega}/2\tilde{Q}$  [38], and the Purcell factor is the enhancement at the peak. This relationship between  $Q$  and LDOS bandwidth is the key to obtaining a tractable well-posed optimization problem in the next section.

We propose that LDOS<sub>*j*</sub> or its variants can be used as figure of merit for the characterization and optimization of a microcavity. Note that the precise figure of merit should really depend on the application. For example, if we are interested in the spontaneous emission rate for the dipole at a specific position  $\mathbf{x}'$  with a specific polarization  $\hat{\mathbf{e}}_j$ , then LDOS<sub>*j*</sub>( $\omega, \mathbf{x}'$ ) is the most relevant figure of merit. On the other hand, if we have a dipole at a specific point  $\mathbf{x}'$  with a randomly distributed polarization, then the figure of merit would be  $\sum_j \text{LDOS}_j(\omega, \mathbf{x}')$  (for the average case) or  $\min_j \text{LDOS}_j(\omega, \mathbf{x}')$  (for the worst case). In nonlinear devices for frequency conversion (in particular, second harmonic generation), a more relevant figure of merit might be  $\min_n \text{LDOS}_j(\omega_n, \mathbf{x}')$ , where  $\omega_n$  are different frequencies of interest. Instead of at a single point  $\mathbf{x}'$ , if the dipoles of interest are distributed [with probability density function  $s(\mathbf{x})$ ] in a region  $V$  with polarization  $\hat{\mathbf{e}}_j$ , the most relevant figure of merit in this case is  $\int_V \text{LDOS}_j(\omega, \mathbf{x}) s(\mathbf{x}) d\mathbf{x}$ . Depending on the applications for enhancement or inhibition, we should maximize or minimize the figure of merit correspondingly. Many other variations could be devised.

For all the above mentioned applications, LDOS<sub>*j*</sub> is the basic building block. We therefore focus on this case: we maximize the spontaneous emission rate for a dipole at a point  $\mathbf{x}'$  with given polarization  $\hat{\mathbf{e}}_j$ . For simplicity, from now on, we shall omit the explicit  $j$  and  $\mathbf{x}'$  dependence from LDOS<sub>*j*</sub>( $\omega, \mathbf{x}'$ ), and simply write LDOS( $\omega$ ) to denote the figure of merit given in Eq. (9). In Sec. 8.2, however, we will also consider a case where the dipole polarization is randomly distributed, and in Secs. 8.2.3 and 8.3 we consider multi-mode cavities.

#### 4. Frequency-averaged LDOS

In previous section, we proposed that one way of attacking the problem of microcavity design is to maximize the LDOS. However, simply maximizing LDOS( $\omega$ ) in Eq. (9) is still ill-posed

as in Sec 2, being equivalent to  $Q/V$ . As explained in Sec. 2, a well-posed formulation could be obtained by specifying a desired cavity lifetime  $\tilde{Q}$ . In terms of LDOS, this is equivalent to specifying an upper bound  $\tilde{\omega}/2\tilde{Q}$  on the bandwidth  $\tilde{\Gamma}$ . For computational convenience, we instead maximize the *average* LDOS over this finite bandwidth  $\tilde{\Gamma}$ :

$$L = \int_{-\infty}^{\infty} \text{LDOS}(\omega)W(\omega)d\omega. \quad (11)$$

Here,  $W(\omega)$  is some weight function or window function we choose, which is peaked around the design frequency  $\tilde{\omega}$  and decays rapidly (with a finite integral) outside of a bandwidth  $\tilde{\Gamma}$  around  $\tilde{\omega}$ . As we will explain in Sec. 4.3, it will turn out that performing this average is mathematically equivalent to computing the LDOS at a single frequency with an absorption loss added into the system, which effectively limits  $Q$  to the desired bandwidth (more precisely, for  $Q \gg \tilde{Q}$  the figure of merit is dominated by the modal volume). Equivalently, as  $Q$  increases the LDOS of a resonance approaches a delta function with amplitude  $\sim 1/V$  [38], so that the integral Eq. (11) is determined mainly by the  $1/V$  factor as soon as the resonance is narrower than  $\tilde{\Gamma}$  ( $Q \gg \tilde{Q}$ ). The key point is that maximizing the bandwidth-averaged LDOS regularizes the optimization problem to eliminate the possibility of diverging- $Q$  solutions as the computational cell size increases.

The main remaining question is how to efficiently compute the average LDOS of Eq. (11). The most straightforward approach would be to apply a standard numerical-integration procedure [43], which would involve evaluating the LDOS (i.e., solving a scattering problem) for many separate frequencies over the bandwidth  $\tilde{\Gamma}$ . (This is similar in spirit to the multi-frequency optimization approach of [24].) However, as described in the next two sections, we can exploit techniques from complex analysis to evaluate the *exact* LDOS integral by solving only a *single* scattering problem at a *complex* frequency  $\tilde{\omega} + i\tilde{\Gamma}$ . The key to this technique is that the LDOS derives from a causal Green's function, as reviewed in Sec. 4.1. This allows us to perform a contour integration, with an appropriate choice of  $W(\omega)$ , in order to obtain the integral as described in Sec. 4.2. As explained in Sec. 4.3, we can reinterpret any complex-frequency scattering problem as a real-frequency scattering problem with complex materials (i.e., absorption). Finally, in Sec 5 we discuss convenient choices of the window function  $W(\omega)$  that reduce the contour integral to a *single* complex- $\omega$  scattering problem. (We previously applied a very similar application of causality and complex analysis to analysis of electromagnetic cloaking bandwidth [11], and related ideas can be found in quantum field theory [44, 45].)

#### 4.1. Causality and analyticity

Before we proceed, we first define a function  $f(\omega)$ , which is a complex version of  $\text{LDOS}(\omega)$ :

$$f(\omega, \mathbf{x}') = -\frac{6}{\pi} \int \mathbf{J}^*(\mathbf{x}) \cdot \mathbf{E}(\mathbf{x}, \omega) d\mathbf{x}. \quad (12)$$

Comparing with Eq. (9), it is clear that  $\text{LDOS}(\omega) = \text{Re}[f(\omega)]$ . [From now on, we will omit the  $\mathbf{x}'$  dependency of  $f(\omega, \mathbf{x}')$ .] Note that the operator  $\mathcal{M}(\epsilon, \omega)$  given in Eq. (8) is a linear operator relating the electric field  $\mathbf{E}(\mathbf{x}, \omega)$  to the (time-harmonic) input electric current  $\mathbf{J}(\mathbf{x})$  at a given frequency  $\omega$ . Causality [the electric field  $\mathbf{E}$  comes after (not before) the current  $\mathbf{J}$ ] implies that  $\mathbf{E}(\mathbf{x}, \omega)$  is analytic in the upper-half complex- $\omega$  plane [46]. Therefore,  $f(\omega)$  is also analytic in the upper-half complex- $\omega$  plane.

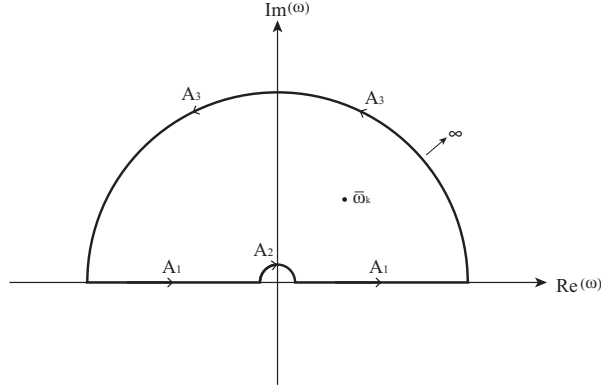


Fig. 2. Contour integration path. The frequency-averaged LDOS is the path integral along arc  $A_1$  in the limit of an infinite-radius arc. By choosing the proper window/weight function  $W(\omega)$  for optimizing LDOS in a desired bandwidth, the contribution along arcs  $A_2$  and  $A_3$  can be made negligible compared to  $A_1$ . Therefore, the residues at poles  $\bar{\omega}_k$  enclosed by this contour can be used to approximate the averaged LDOS.

#### 4.2. Contour integration

In this section, we are going to compute the mean LDOS by exploiting the analyticity of  $f(\omega)$ , via:

$$L = \int_{-\infty}^{\infty} \text{LDOS}(\omega)W(\omega)d\omega = \int_{-\infty}^{\infty} \text{Re}[f(\omega)]W(\omega)d\omega = \text{Re} \left[ \text{p.v.} \int_{-\infty}^{\infty} f(\omega)W(\omega)d\omega \right]. \quad (13)$$

Here, p.v. denotes the Cauchy principal value, which we use because the imaginary part of  $f(\omega)$  may have a singularity at  $\omega = 0$ , as in the case of  $f(\omega)$  in free space [40,42,47,48]. Now we want to complete our integration contour (Fig. 2) in the upper-half plane and evaluate  $L$  by residue theorem [49]

$$\int_{A_1} + \int_{A_2} + \int_{A_3} f(\omega)W(\omega)d\omega = 2\pi i \sum_k \text{Res}[f(\omega)W(\omega), \bar{\omega}_k]. \quad (14)$$

Here,  $\bar{\omega}_k$  denotes a pole of  $W(\omega)$  in the upper-half plane and  $\text{Res}$  is its residue [49]. In Sec. 5, we will choose  $W(\omega)$  so that these poles and residues are easy to evaluate. Furthermore, we can choose the weight function  $W(\omega)$  so that it decays faster than  $1/|\omega|^3$  for large  $\omega$ , in which case the contribution from arc  $A_3$  will be zero since  $\text{LDOS}(\omega)$  is proportional to  $\omega^2$  in 3d (or  $\omega$  in 2d) free space and  $f(\omega)W(\omega)$  will decay faster than  $1/|\omega|$  on arc  $A_3$ . We can obtain the contribution from arc  $A_2$  by evaluating the residue due to the simple pole of  $f(\omega)$  at  $\omega = 0$

$$\int_{A_2} f(\omega)W(\omega)d\omega = -\frac{1}{2}2\pi i W(0) \lim_{\omega \rightarrow 0} \omega f(\omega). \quad (15)$$

The factor  $-1/2$  comes from the fact that the integration is along a clockwise semicircle. Since the weight function is peaked around design frequency  $\bar{\omega}$  with some narrow bandwidth  $\tilde{\Gamma}$ , we can require that  $W(0)$  is small, and therefore the contribution from  $A_2$  is negligible comparing to the residues at  $\bar{\omega}_k$ . From Eq. (13),  $L$  is just the path integral along the path  $A_1$  [46]. Therefore,

we have

$$\begin{aligned}
L &= \text{Re} \left[ 2\pi i \sum_k \text{Res} [f(\omega)W(\omega), \bar{\omega}_k] - \left( \int_{A_2} + \int_{A_3} f(\omega)W(\omega)d\omega \right) \right] \\
&\approx \text{Re} \left[ 2\pi i \sum_k \text{Res} [f(\omega)W(\omega), \bar{\omega}_k] \right].
\end{aligned} \tag{16}$$

#### 4.3. From complex frequency to material absorption

To compute the residue at the complex poles, we need to solve the scattering problem at complex frequencies. More precisely, the scattering problem Eq. (8) at complex frequency  $\omega + i\Gamma$  can be written as

$$\begin{aligned}
&\left( \nabla \times \frac{1}{\mu(\mathbf{x})} \nabla \times -\varepsilon(\mathbf{x})(\omega + i\Gamma)^2 \right) \mathbf{E}(\mathbf{x}, \omega + i\Gamma) = i(\omega + i\Gamma)\mathbf{J}(\mathbf{x}) \\
\iff &\left( \nabla \times \frac{1}{\mu(\mathbf{x})(1 + i\frac{\Gamma}{\omega})} \nabla \times -\varepsilon(\mathbf{x})\omega^2 \left( 1 + i\frac{\Gamma}{\omega} \right) \right) \mathbf{E}(\mathbf{x}, \omega + i\Gamma) = i\omega\mathbf{J}(\mathbf{x}) \\
\iff &\left( \nabla \times \frac{1}{\mu(\mathbf{x})(1 + \frac{i}{2Q})} \nabla \times -\varepsilon(\mathbf{x}) \left( 1 + \frac{i}{2Q} \right) \omega^2 \right) \mathbf{E}(\mathbf{x}, \omega + i\Gamma) = i\omega\mathbf{J}(\mathbf{x}).
\end{aligned} \tag{17}$$

We denote this complex scattering operator by  $\tilde{\mathcal{M}}(\varepsilon, \omega)$ , namely

$$\tilde{\mathcal{M}}(\varepsilon, \omega) = \nabla \times \frac{1}{\mu(\mathbf{x})(1 + \frac{i}{2Q})} \nabla \times -\varepsilon(\mathbf{x}) \left( 1 + \frac{i}{2Q} \right) \omega^2 = \nabla \times \frac{1}{\tilde{\mu}(\mathbf{x})} \nabla \times -\tilde{\varepsilon}(\mathbf{x})\omega^2. \tag{18}$$

Clearly, this is equivalent to solving a scattering problem at real frequency  $\tilde{\omega}$  with materials  $\tilde{\varepsilon}(\mathbf{x}) = \varepsilon(\mathbf{x})(1 + i/2Q)$  and  $\tilde{\mu}(\mathbf{x}) = \mu(\mathbf{x})(1 + i/2Q)$ . (In fact, any change to the frequency can be converted into a change of materials [50].) In particular, adding a positive imaginary part to  $\omega$  (for  $\bar{\omega}_k$  in the upper-half plane [11]) corresponds to a positive imaginary part in  $\tilde{\varepsilon}(\mathbf{x})$  and  $\tilde{\mu}(\mathbf{x})$ , which corresponds (with our  $e^{-i\omega t}$  convention) to an absorption loss.

## 5. Possible window functions

In this section, we discuss two convenient window functions: a simple Lorentzian and the square of a Lorentzian. (A third possibility, the difference of two Lorentzians is discussed in [51].)

### 5.1. A simple Lorentzian

The simplest window function, with only a simple pole in the upper-half plane, is a Lorentzian centered at  $\tilde{\omega}$  with half-width  $\tilde{\Gamma}$ . The frequency-average LDOS against this weight is

$$L_1 = \int \text{LDOS}(\omega) \frac{\tilde{\Gamma}/\pi}{(\omega - \tilde{\omega})^2 + \tilde{\Gamma}^2} = \text{Re}[f(\tilde{\omega} + i\tilde{\Gamma})], \tag{19}$$

which only requires solving the scattering problem Eq. (17) at a single complex-frequency  $\tilde{\omega} + i\tilde{\Gamma}$ .  $L_1$  is a perfectly finite, well-defined quantity in a discretized simulation with a finite spatial resolution (finite grid).

In combination with Sec. 4.3, the objective Eq. (19) has a simple interpretation that coincides with the discussion in Sec. 2. Our frequency-averaged LDOS objective is equivalent to

maximizing the LDOS at a single frequency, i.e. the Purcell factor  $Q/V$  of a cavity, but a cavity in which an absorption loss has been added to otherwise lossless materials. In particular, the absorption loss that arises from multiplying  $\epsilon(\mathbf{x})$  and  $\mu(\mathbf{x})$  by  $1 + i/2\tilde{Q}$  can be seen from perturbation theory [14] (for  $\tilde{Q} \gg 1$ ) to yield an absorption lifetime of exactly  $Q_{\text{absorption}} = \tilde{Q}$ . From Eq. (5), this means that  $Q_{\text{total}} \leq \tilde{Q}$ , and that increasing  $Q_{\text{rad}} \gg \tilde{Q}$  will have little effect on the Purcell factor  $Q_{\text{tot}}/V$ . Hence, optimizing the frequency-averaged objective will tend to increase  $Q_{\text{rad}}$  until it is  $> \tilde{Q}$ , and after that will mainly try to decrease  $V$ , similar in spirit to Eq. (6).

However, a careful examination reveals that this simple average does not converge as the resolution increases. There are two equivalent ways to understand this. First, in a continuous medium, the integral does not converge because the window function decays like  $1/|\omega|^2$  while  $\text{LDOS}(\omega)$  behaves like  $|\omega|$  (in 2d free space) or  $|\omega|^2$  (in 3d free space) for large  $|\omega|$ . (For finite spatial resolution, there is an upper frequency cutoff that eliminates this divergence.) Second, from the relationship between the complex-frequency scattering and lossy material discussed in the previous section, we know that the residue  $\text{Re}[f(\tilde{\omega} + i\tilde{\Gamma})]$  is actually the power emitted by a dipole in lossy material, which is the sum of the power radiating to the outside of the cavity and the power absorbed by the lossy material in the cavity [41]. It is known that this absorbed power is infinite because  $\mathbf{E}(\mathbf{x})$  diverges as  $1/r^3$  in the neighborhood of the dipole (in 3d free space) [40, 47, 52, 53]. (In a lossless medium, only  $\text{Im}[\mathbf{E}(\mathbf{x})]$  diverges as  $1/r^3$ , so  $\text{LDOS} \sim \text{Re}[\mathbf{E}(\mathbf{x})]$  is finite.) In discretized space, the Green's function is finite and diverges as (resolution)<sup>3</sup> in 3d. To avoid this singularity, we need to choose window functions which decay faster than  $|\omega|^3$  at large  $|\omega|$ . Two natural candidates are the difference of two Lorentzians [51] and the square of a Lorentzian.

## 5.2. Square of a Lorentzian

To ensure that the  $W(\omega)$  decays faster than  $1/|\omega|^3$ , we propose the window function,

$$W(\omega) = \frac{2\tilde{\Gamma}^3/\pi}{((\omega - \tilde{\omega})^2 + \tilde{\Gamma}^2)^2}, \quad (20)$$

which is a normalized square of a Lorentzian function. This window function has a double pole at  $\omega = \tilde{\omega} + i\tilde{\Gamma}$  in the upper-half plane. Applying the residue theorem, we have from Eq. (16) and Eq. (20)

$$L = \int_{-\infty}^{\infty} \text{LDOS}(\omega)W(\omega)d\omega \approx \text{Re} [f(\tilde{\omega} + i\tilde{\Gamma}) - i\tilde{\Gamma}f'(\tilde{\omega} + i\tilde{\Gamma})], \quad (21)$$

where  $f'(\cdot)$  denotes differentiation with respect to  $\omega$ . In appendix 10, we show that

$$f'(\omega, \mathbf{x}') = \frac{f(\omega, \mathbf{x}')}{\omega} + i\frac{12}{\pi} \int \epsilon(\mathbf{x})\mathbf{E}^T(\mathbf{x}, \omega)\mathbf{E}(\mathbf{x}, \omega)d\mathbf{x}. \quad (22)$$

From Eq. (12) and Eq. (22), it is clear that both  $f(\tilde{\omega} + i\tilde{\Gamma})$  and  $f'(\tilde{\omega} + i\tilde{\Gamma})$  can be obtained from a single scattering solution  $\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})$  (see appendix 10) and

$$\begin{aligned} & f(\tilde{\omega} + i\tilde{\Gamma}) - i\tilde{\Gamma}f'(\tilde{\omega} + i\tilde{\Gamma}) \\ &= \frac{\tilde{\omega}}{\tilde{\omega} + i\tilde{\Gamma}} \left(-\frac{6}{\pi}\right) \hat{\mathbf{e}}_j^* \cdot \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) + \frac{12}{\pi} \tilde{\Gamma} \int \epsilon(\mathbf{x})\mathbf{E}^T(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x}. \end{aligned} \quad (23)$$

In summary, we can still obtain the entire frequency-averaged LDOS by solving a single scattering problem Eq. (17) at a complex frequency  $\tilde{\omega} + i\tilde{\Gamma}$ .



We know that Eq. (20) gives a finite average LDOS because it decays fast enough with  $\omega$ , but it is interesting to also consider how it fixes the divergence from the second viewpoint in Sec. 5.1 (that of the infinite power absorption from a dipole in a lossy medium). The explanation is essentially that the second term in Eq. (23) is roughly a subtraction of the divergent absorbed power from the first term:  $\tilde{\Gamma}\varepsilon$  is  $\omega \text{Im}(\tilde{\varepsilon})$  from Sec. 5.1, and  $\omega \text{Im}(\tilde{\varepsilon})|\mathbf{E}|^2$  is absorbed power [41]. (A subtlety arises from having  $+\mathbf{E}^T\mathbf{E}$ , rather than  $-|\mathbf{E}|^2$ , but the  $1/r^3$  divergence at  $r \rightarrow 0$  should be dominant in  $\text{Im}(\mathbf{E})$  for small  $\tilde{\Gamma}$  so one should have  $\mathbf{E}^T\mathbf{E} \approx -\text{Im}(\mathbf{E}^T)\text{Im}(\mathbf{E}) \approx -|\mathbf{E}|^2$  as  $r \rightarrow 0$ .)

Since the role of the second term in Eq. (23) is essentially to subtract off the divergent absorbed power in lossy  $\tilde{\varepsilon}$  medium, and this divergence comes from the  $1/r^3$  field divergence that is independent of geometry (the scattered field from the surrounding geometry is finite at  $r = 0$ ), one might expect that the second term in Eq. (23) plays little role in geometry optimization at a fixed resolution. Indeed, we find in numerical experiments that the optimizations *with* and *without* the second terms in Eq. (23) for the 2d TE case (discussed in Sec. 8.2) discover similar structures. Therefore, in Sec. 8 we optimize the simpler single-Lorentzian objective of section 5.1, although Eq. (23) is computationally feasible if it becomes necessary.

## 6. A preliminary formulation

Now we have a preliminary formulation for our cavity optimization in terms of the frequency-averaged local density of states:

$$\max_{\{\text{designs}\}} L = \int_{-\infty}^{\infty} \text{LDOS}(\omega)W(\omega)d\omega. \quad (24)$$

We can evaluate the objective  $L$  by contour integration, which only requires us to solve the complex scattering problem Eq. (17) once. If we choose the window function  $W(\omega)$  from Eq. (20), then the problem can be reformulated as

$$\max_{\{\text{designs}\}} L = \text{Re} [\mathcal{L} = f(\tilde{\omega} + i\tilde{\Gamma}) - i\tilde{\Gamma}f'(\tilde{\omega} + i\tilde{\Gamma})]. \quad (25)$$

Mathematically, to compute the objective in Eq. (25), we need

1. For given  $\varepsilon(\mathbf{x})$ ,  $\tilde{\omega}$ , and  $\tilde{\Gamma}$ , solve the complex scattering problem Eq. (17) to obtain  $\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})$ .
2. Obtain  $\mathcal{L}$  from the solution  $\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})$  through the formula Eq. (23).
3. Take the real part of  $\mathcal{L}$  to get  $L$ .

To speed up the optimization, we must also have the gradient of the objective with respect to the design parameter  $\varepsilon_k$  (the dielectric constant at each ‘‘pixel’’  $k$ ). Applying the standard adjoint technique [54], only one more linear system with the same operator  $\mathcal{M}(\varepsilon, \tilde{\omega})$  but a different source term need be solved to obtain the gradient. We provide the detailed calculation in appendix 10 and summarize the procedure here:

1. Solve the complex scattering problem

$$\mathcal{M}(\varepsilon, \tilde{\omega})\mathbf{A}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) = \varepsilon(\mathbf{x})\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) \quad (26)$$

to obtain  $\mathbf{A}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})$ .

2. The gradient  $\partial\mathcal{L}/\partial\epsilon_k$  is a combination of  $\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})$  and  $\mathbf{A}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})$  (see appendix 10):

$$\frac{\partial\mathcal{L}}{\partial\epsilon_k} = \left(i + \frac{1}{\tilde{Q}}\right) \frac{2}{\pi} \tilde{\omega} \mathbf{E}^T(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \mathbf{E}(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) + \frac{12}{\pi} \frac{\tilde{\omega}^3}{\tilde{Q}} \left(1 + \frac{i}{2\tilde{Q}}\right) \mathbf{A}^T(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \mathbf{E}(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}). \quad (27)$$

3. Take the real part of  $\partial\mathcal{L}/\partial\epsilon_k$  to obtain  $\partial L/\partial\epsilon_k$ .

Note that the scattering operator Eq. (26) in the sensitivity analysis is the same as the operator in the objective evaluation. We can take advantage of this by reusing the information (e.g., the preconditioner or LU factorization) from the solution of Eq. (17). We will discuss this in detail in Sec. 7.1.

## 7. Numerical scheme for cavity optimization

In this section, we discuss the numerical implementation for our frequency-averaged LDOS formulation given in Sec. 6. In order to solve this PDE-constrained optimization problem computationally, we need fast and efficient implementations for objective evaluation, gradient evaluation, and optimization.

### 7.1. Objective and gradient evaluation

As we discussed in Sec. 6, evaluating the objective LDOS involves solving the scattering problem Eq. (17). We can apply any standard frequency-domain solver technique to this problem (e.g., finite-difference, finite-element, or boundary-element methods). Here, we simply adopt the finite-difference approach [55–57]. If we impose mirror symmetry planes in the system, we can obtain an 8-fold reduction in the number of unknowns (see Sec 8.5.1).

For the finite-difference frequency-domain (FDFD) method, the most robust solution technique is a sparse-direct solver, which is excellent in 2d, but expensive (in both memory and time) in 3d [58]. In contrast to direct solvers, iterative solvers (e.g., GMRES or BiCGStab) work quite well if one has a good preconditioner [36]. Here, we combine both of these techniques. During the optimization, we re-solve many times for slightly different structures. Therefore, we can use sparse-direct factorization from one step as a preconditioner for iterative solvers in many subsequent steps [51]. We implemented the FDFD solver with the parallel sparse-matrix library PETSc [59–61] and the parallel sparse-direct solver PaStiX [62].

### 7.2. Optimization scheme

We use a free-software implementation [63] of standard gradient-based optimization algorithms, and we find that low-storage BFGS [64] and conservative convex separable approximations (CCSA) [65] work equally well. However, we find that one additional mathematical transformation is required in order to optimize high- $Q$  cavities: instead of maximizing  $L$ , we minimize  $1/L$ .

If we attempt to maximize  $L$  directly as in Eq. (24), we typically find that the convergence of any standard optimization algorithm slows to a halt for  $Q \gtrsim 1000$ . The reason for this is that, for high- $Q$  cavities, the objective function becomes a “narrow ridge” (a sharply-peaked function along some low-dimensional manifold in the parameter space), with a large second derivative ( $\sim Q^3$ , as shown in [51]) in the direction perpendicular to the ridge, and it is well known that optimization along narrow ridges is problematic [66]. Most algorithms tend to “zig-zag” slowly along the ridge, and even quasi-Newton methods like BFGS break down when the

second derivative is so large that the Hessian matrix becomes ill-conditioned. However, because the LDOS is strictly positive, there is a simple solution: maximizing  $L$  is equivalent to minimizing  $1/L$ , and the reciprocal of a sharp peak is a shallow valley, so we find that the  $1/L$  objective avoids most of the problems of slow convergence. (For example, in the 2d TM optimization to be discussed in Sec. 8.1, maximization of  $L$  makes no progress if the initial structure is a high- $Q$  photonic-crystal cavity, whereas minimization of  $1/L$  converges to a significantly improved structure as described below.)

In practice, there is another useful technique: a successive-refinement strategy (somewhat analogous to [67, 68]). We found that gradually increasing the specified  $\tilde{Q}$  (decreasing the bandwidth  $1/\tilde{Q}$ ) tends to reduce the likelihood that the optimization becomes trapped in a poor local optimum close to the starting structure, and usually produces a much better result at the highest  $\tilde{Q}$ . That is, we optimize for  $\tilde{Q} = 10$ , then use that result as a starting point for optimization at  $\tilde{Q} = 100$ , and so on.

## 8. Results for cavity optimization

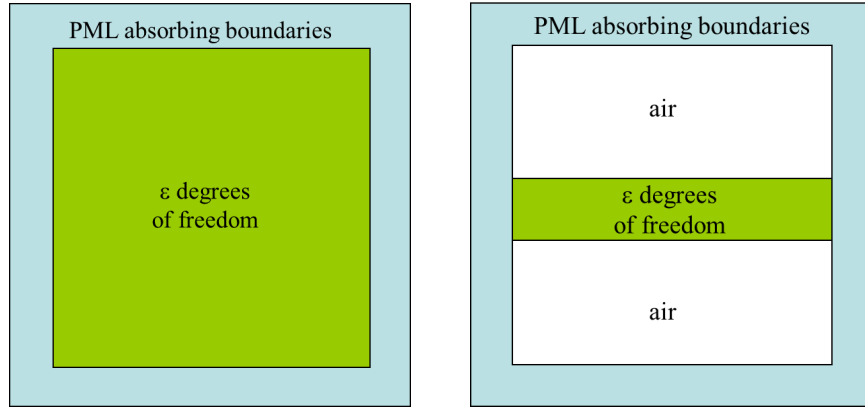
In this section, we present some 2d and 3d cavity-optimization results, summarized as follows. We start with high-resolution 2d cases, and run simulations with different initial guesses (vacuum, photonic crystal with a defect, and random structures) and different dipole polarizations (TM, TE, and random). In the region to be optimized, we allow the dielectric constant  $\epsilon \in [1, 12.4]$  at each pixel to be one degree of freedom [Fig. 3(a)]. For the 2d TE case, the optimization discovers similar structures for maximizing the spontaneous emission rate of a *specific* dipole polarization and a *randomly* polarized dipole. However, optimizing the *worst* case of a randomly polarized dipole finds a three-fold symmetric structure (Sec. 8.2.3), while two-frequency cavity optimizations yield more complex patterns depending on the polarization (Sec. 8.3). In another 2d scenario, to obtain the  $Q$  versus  $V$  tradeoff analogous to 3d slabs, we limit the degrees of freedom in one direction and choose a thin strip, instead of a square, as the region for optimization [Fig. 3(b)]. As the degrees of freedom increase, the radiation  $Q$  first increases and then becomes saturated, limited by the numerical precision in the computation. Finally, we ran a full 3d optimization on a supercomputer and obtained a structure with extremely small mode volume  $V = 0.06(\lambda/n)^3$ , ( $n = \sqrt{12.4}$ ). In the following, we are minimizing  $1/L$  (the inverse of the frequency-averaged LDOS) and variations thereof, as described in the previous sections, but for simplicity we describe this below as “maximizing LDOS.”

Here, the focus is on illustrating the formulation and on establishing theoretical bounds. Further regularization techniques are generally required to obtain easily manufacturable structures, as discussed in Sec. 9.

### 8.1. 2D TM case

In this section, we maximize the LDOS for a dipole with TM polarization ( $\mathbf{E}$  out of plane) in a 2d setting [Fig. 3(a)]. One possible starting point is a photonic crystal with a defect [14], like the one shown in Fig. 4(a). This is a periodic arrangement (periodicity  $a$ ) of dielectric silicon rods (radius  $0.2a$  and permittivity  $\epsilon = 12.4$ ) with one defect rod at the center (radius  $0.1a$ ). The defect TM mode is at frequency  $0.32(2\pi/a)$ , with quality factor  $Q = 1.41 \times 10^8$  and mode volume  $V = 0.097(\lambda/n)^2$ . With this structure as an initial guess, we run the optimization and obtain an entirely different nested-ring structure [Fig. 4(b)] with quality factor  $Q = 1.01 \times 10^{10}$  and mode volume  $V = 0.075(\lambda/n)^2$ . Clearly, the optimization itself discovers a “radially periodic” structure, reminiscent of a Bragg onion [69]. We also run the optimization with vacuum as initial guess and obtain similar structure (Fig. 5) with  $Q = 1.30 \times 10^9$  and  $V = 0.075(\lambda/n)^2$ .

In these two optimizations, we gradually increase the specified  $\tilde{Q}$  (we decrease the bandwidth



(a). A square region for degrees of freedom.

(b). A thin strip region for degrees of freedom.

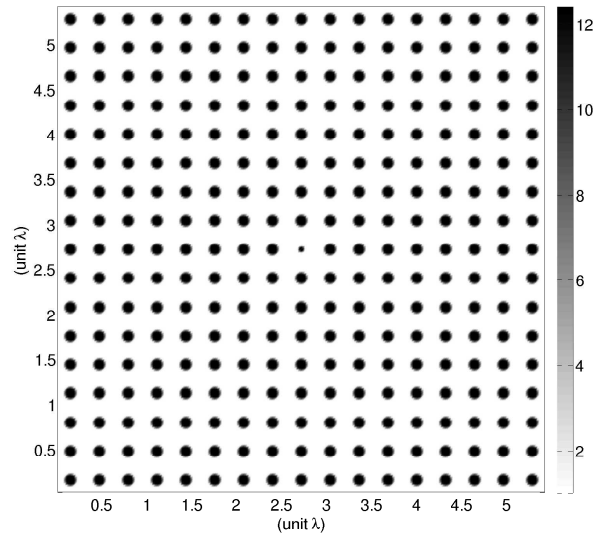
Fig. 3. For 2d cavity optimization, we start in Secs. 8.1–8.3 by optimizing over every pixel in the interior of the computational domain as indicated in (a). This leads to cavities that utilize bandgap structures to confine light with arbitrary  $Q$ , regardless of  $V$ , limited only by the size of the domain. In order to investigate  $Q$  vs.  $V$  tradeoffs analogous to those in 3d slabs, in Sec. 8.4 we limit the degrees of freedom to a thin strip (b), which imposes intrinsic radiation losses (perpendicular to the strip) and forces the optimization to sacrifice  $V$  in order to increase  $Q$ . A full 3d optimization is considered in Sec. 8.5.

$1/\tilde{Q}$  from 10 to  $10^5$ . The optimization at  $\tilde{Q} = 10$  actually gives a high  $Q$  cavity (almost the same radiation  $Q$ ) with the resonance at about  $1.003\tilde{\omega}$ . The optimizations at higher  $\tilde{Q}$  simply tune this structure so that the resonant frequency becomes much closer to  $\tilde{\omega}$ . In a two-dimensional situation such as this, there is no intrinsic  $Q$  versus  $V$  tradeoff, unlike in 3d slabs [14], because  $Q \rightarrow \infty$  for a finite modal volume  $V$  as the number of layers in a Bragg onion increases [69]. So, the optimization in such a 2d case is mainly minimizing  $V$ , while the  $Q$  is bounded only by the computational-cell size.

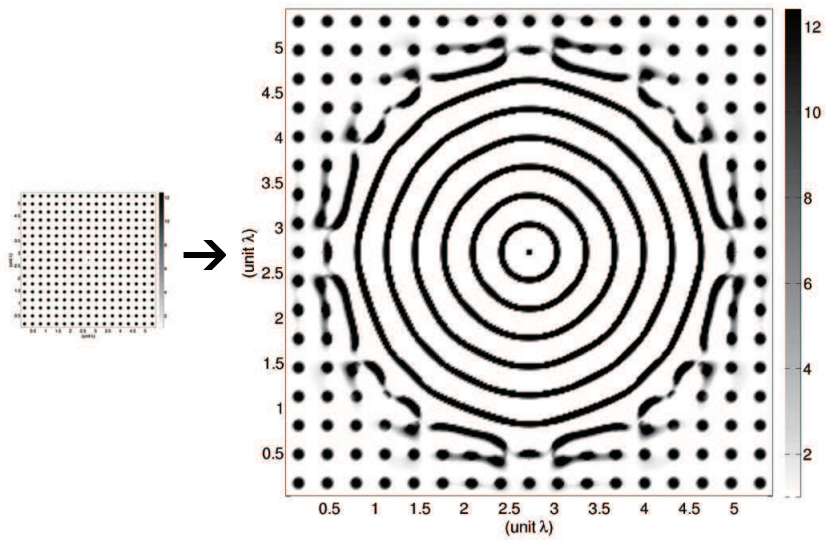
Note that we allow the dielectric permittivity of each pixel to vary continuously from  $\epsilon_{\min} = 1.0$  to  $\epsilon_{\max} = 12.4$ , but almost all the pixels (except a few at the interfaces) are at either  $\epsilon_{\min}$  or  $\epsilon_{\max}$  in the optimized structure. This phenomenon (reminiscent of “bang-bang” solutions in control theory [70]) had also been observed empirically in other cavity-related optimization work [3, 71, 72]. There has been some recent progress in proving theoretically that this is the expected solution: Osting and Weinstein [73] recently analyzed optimization problems for scalar waves, and showed that maximizing an energy confinement time over the permittivity at every point in space generally leads to a solution in which the permittivity is either the maximum or the minimum allowed value at every point, excepting a set of measure zero (e.g., at the interfaces between regions) in the limit as the resolution goes to infinity. However, we also obtain counterexamples for other types of cavity optimization, e.g. in the doubly-degenerate cavity design of Sec. 8.2.3, where substantial regions converge to intermediate values. Fortunately, in cases where this occurs, there are a variety of techniques to obtain binary solutions as discussed in Sec. 9.

## 8.2. 2D TE case

In this section, we consider the 2d TE polarization, for which (unlike the TM polarization of Sec. 8.1) the objective function breaks rotational symmetry and we do not expect similar Bragg-onion solutions. We will start with maximizing LDOS for a fixed  $\hat{\mathbf{e}}_x$  polarization in



(a). PhC cavity initial guess for 2d TM optimization. It has quality factor  $Q=1.41 \times 10^8$  and mode volume  $V = 0.097(\lambda/n)^2$ .



(b). 2d TM optimized structure with  $Q=1.01 \times 10^{10}$  and  $V = 0.075(\lambda/n)^2$  obtained from PhC initial guess.

Fig. 4. 2d TM optimization from PhC cavity initial guess (Sec. 8.1).



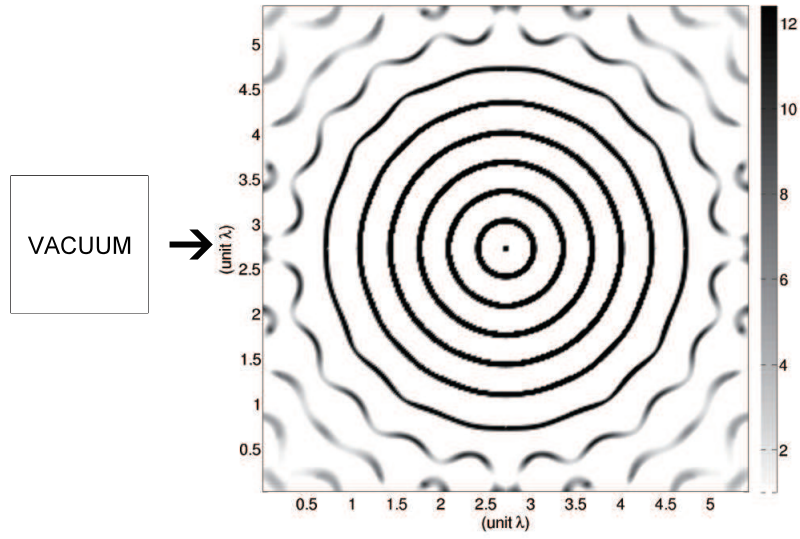


Fig. 5. 2d TM optimization from vacuum initial guess (Sec. 8.1).  $Q=1.30\times 10^9$  and  $V = 0.075(\lambda/n)^2$ .

8.2.1. For a randomly polarized dipole, we consider two cases: maximizing the average over all polarizations in 8.2.2 and maximizing the minimum (worst case) over all polarizations in 8.2.3.

8.2.1. Fixed dipole polarization:  $\max \text{LDOS}(\tilde{\omega}; \hat{\mathbf{e}}_x)$

For the 2d TE polarization, let us first look at the case where the dipole is polarized in the  $\hat{\mathbf{e}}_x$  direction. In other words, we want to maximize  $\text{LDOS}(\tilde{\omega}; \hat{\mathbf{e}}_x)$ . From a vacuum initial guess, the optimization discovers the structure shown in Fig. 6. This structure has quality factor  $Q=5.16\times 10^8$  and mode volume  $V = 0.092(\lambda/n)^2$ . [Again, the  $\tilde{Q} = 10$  gives an equally high- $Q$  cavity with resonant frequency at  $1.0007\tilde{\omega}$ . The optimizations at higher  $\tilde{Q} = 10^2$  to  $10^5$  simply tune the resonant frequency to  $\tilde{\omega}$ .] For a random initial guess (uniform random pixels), we also obtain the same structure as the one from a vacuum initial guess, which suggests that the result may be close to a global optimum.

8.2.2. Randomly polarized dipole:  $\max \text{mean}_j \text{LDOS}(\tilde{\omega}; \hat{\mathbf{e}}_j)$

Now we want to study the case in which the dipole is randomly polarized in the plane and the objective is the mean LDOS. One might hope that the optimization will find a structure that resonates for both polarizations at the same frequency, and hence by linearity will resonate for all in-plane polarizations—this corresponds to the requirement that the microcavity be doubly degenerate, and many symmetry groups besides circular symmetry can support double degeneracies (such as three-fold, four-fold, or six-fold symmetrical structures [74]). However, we find that this is not the case: the optimization finds a singly resonant cavity that enhances LDOS for one polarization (chosen “randomly” depending on the initial structure) at the expense of the other polarization. As explained below, this suggests that the LDOS of the best single-mode TE cavity is more than twice the LDOS of the best doubly degenerate TE cavity. We performed the mean-LDOS optimization as follows. It is easy to show [29, problem 8.6] that maximizing the LDOS for a random polarization by averaging *all* polarizations is equivalent to maximizing

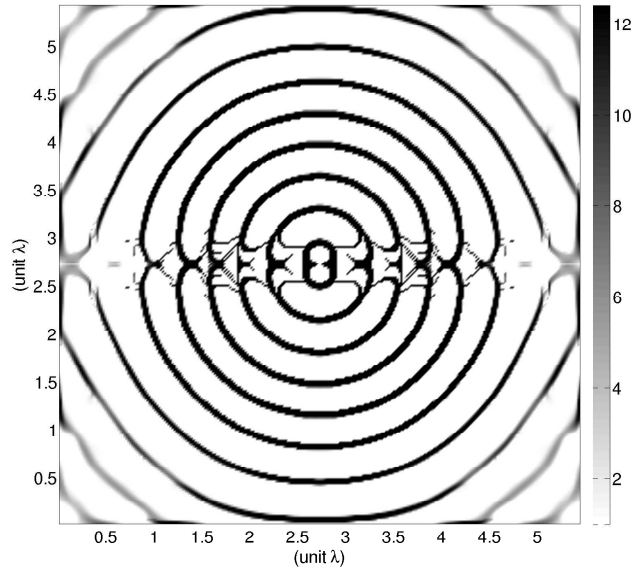


Fig. 6. 2d TE optimization for  $\hat{\mathbf{e}}_x$  polarization from Sec. 8.2.1. The structure has  $Q=5.16 \times 10^8$  and  $V = 0.092(\lambda/n)^2$  obtained from vacuum initial guess.

the mean from the  $\hat{\mathbf{e}}_x$  and  $\hat{\mathbf{e}}_y$  polarizations, namely  $[\text{LDOS}(\omega; \hat{\mathbf{e}}_x) + \text{LDOS}(\omega; \hat{\mathbf{e}}_y)]/2$ . For this new objective, we ran many different simulations with different pseudo random initial guesses (each pixel is randomly chosen between  $\epsilon_{\min}$  and  $\epsilon_{\max}$ ). About half gave the structure optimizing the  $\hat{\mathbf{e}}_x$  polarization (Fig. 6), while the other half gave the  $90^\circ$ -rotated structure optimizing the  $\hat{\mathbf{e}}_y$  polarization. From these results, it seems that the optimization, instead of favoring both  $\hat{\mathbf{e}}_x$  and  $\hat{\mathbf{e}}_y$  polarization simultaneously, simply randomly picks one direction and optimizes it. That is, there is a *spontaneous symmetry breaking*: it is better to optimize one polarization at the expense of the other than to try to obtain a doubly degenerate cavity that resonates for both polarizations. By selecting only one polarization to enhance, these structures sacrifice a factor of 2 in the mean LDOS, which is why we conclude that the best single-mode LDOS is at least twice as big as the best two-mode LDOS.

### 8.2.3. Optimization for a randomly polarized dipole: $\max \min_j \text{LDOS}(\tilde{\omega}; \hat{\mathbf{e}}_j)$

In the previous section, we showed that maximizing the *mean* LDOS over all in-plane polarizations was equivalent to maximizing the LDOS for a single polarization at the expense of the orthogonal polarization. In order to obtain a structure that enhances all polarizations equally (via a doubly degenerate resonance), we consider a different objective: we maximize the *minimum* LDOS over all in-plane polarizations (rather than the mean). The result of this optimization is a three-fold symmetric microcavity shown in Fig. 7(a), which is the smallest symmetry group that supports a (non-accidental) doubly degenerate mode [74] shown in Fig. 7(b-c). (The rectangular FDFD grid breaks the three-fold symmetry, but the structure converges to true three-fold symmetry with an exact degeneracy as the resolution is increased.) Intuitively, larger symmetry groups have fewer degrees of freedom for optimization (the  $C_{mv}$  symmetry group [74] with  $m$ -fold rotational symmetry plus reflections leaves only a wedge of angle  $\pi/m$  in the degrees of freedom), so the optimization is picking the smallest possible symmetry group in order to

maximize the number of degrees of freedom available for maximizing LDOS.

As predicted in Sec. 8.2.2, the effective modal volume  $V$  of this doubly degenerate cavity is  $0.34(\lambda/n)^2 > 2$  times the volume of the single-mode cavity from Sec. 8.2.1.

Technically, maximizing the minimum LDOS over all polarizations is significantly harder than maximizing the mean LDOS. Unlike the mean LDOS, it is not sufficient to consider only the  $\hat{\mathbf{e}}_x$  and  $\hat{\mathbf{e}}_y$  polarizations, and in principle we must solve for the LDOS at all angles. The simplest approach is to merely sample the set of polarizations to compute the LDOS at many discrete angles, and then to maximize the minimum LDOS over this set. Although we tried sampling up to 20 discrete angles, we found that it was sufficient (obtaining the *same* structure) to sample only three angles (0, 60, and 120). Furthermore, the minimum LDOS over several angles is no longer a differentiable function, but we circumvent that problem by the standard transformation [75] of introducing a “dummy” variable  $t$  and solving  $\max t$  subject to the nonlinear constraints  $t \leq \text{LDOS}_j$  at each angle  $j$ . (The resulting nonlinear-programming problem was solved via the CCSA algorithm [65].)

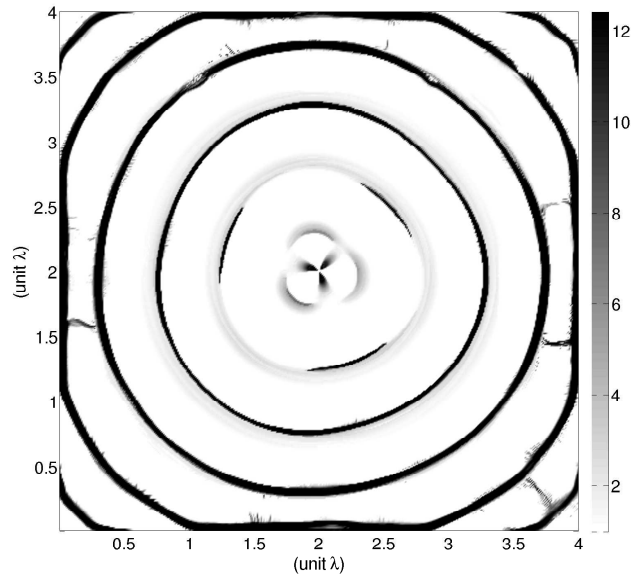
### 8.3. 2D optimization for different frequencies: $\max \min_n \text{LDOS}(\tilde{\omega}_n; \hat{\mathbf{e}}_j)$

In nonlinear devices, e.g. for nonlinear frequency conversion, it is often desirable to design a cavity which resonates at multiple *distinct* frequencies [9, 10, 76–79], leading to a challenging multi-frequency cavity-design problem if one desires a small-volume cavity [80, 81]. The simplest approach to designing such a multi-frequency cavity with our optimization techniques is to maximize the *minimum* LDOS for sources at two (or more) frequencies,  $\tilde{\omega}_1$  and  $\tilde{\omega}_2$ . For example, we begin by considering  $\tilde{\omega}_2 = 2\tilde{\omega}_1$  (coupling TM to TM or TE to TE), the desired relationship for intra-cavity second-harmonic generation (SHG) [79, 81], with results shown in Figs. 8(a) and 8(c). These results exhibit Bragg-onion-like structures similar to the single-frequency optimization in the previous sections, which is not surprising considering that Bragg mirrors tend to have gaps at integer frequency multiples [14] and so the same mirror can confine both a fundamental and harmonic frequency. A more challenging case for optimization, therefore, is to design frequencies that are not integer multiples, and for illustration purposes we considered  $\tilde{\omega}_2 = 1.5\tilde{\omega}_1$ , which results in the more complicated structure shown in Figs. 8(b) and 8(d).

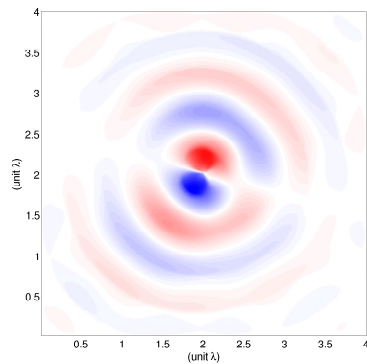
### 8.4. 2D TE thin strip case

In previous 2d TM and TE polarization cases, we expect and obtain almost no  $Q$  versus  $V$  trade off since the cavity can be surrounded by complete photonic bandgap or a Bragg onion, with  $Q$  only limited by the computational-cell size. In a 2d setting, to get a  $Q$  versus  $V$  trade off analogous to 3d slabs, we need to limit the degrees of freedom in one direction in order to force the possibility of radiation loss. In this section, we choose the region for degrees of freedom to be a thin strip [Fig. 3(b)].

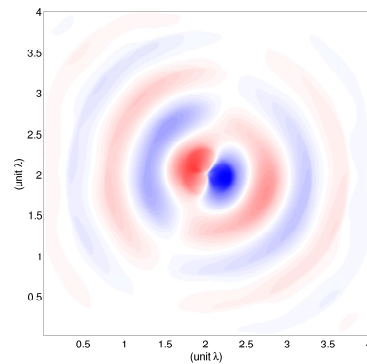
Although we expect cavities in a finite-thickness strip to have intrinsic radiation loss in the direction transverse to the strip, we also expect that it should be possible to make this radiation loss arbitrarily small at the expense of modal volume. For example, if one starts with a periodic waveguide structure [14] and introduces a defect in the periodicity, a resonant mode can be trapped, and the  $Q$  can be made arbitrarily large by gradually tapering from the defect to the periodic structure [14, 82–84]. To explore this tradeoff, we limit the modal volume by considering a finite-length  $d \times 1\lambda$  strip of degrees of freedom. For each value of  $d$  ( $d = 1, 2, 3, 5\lambda$ ), we plot the obtained radiation lifetime  $Q_{\text{rad}}$  as a function of the requested bandwidth  $\tilde{Q}$ , and the result is shown in Fig. 9(a). For any given  $d$ , as  $\tilde{Q}$  is increased then  $Q_{\text{rad}}$  is increased as well (as explained in Sec. 4), until it saturates at a maximum determined by two factors. First, the maximum  $Q_{\text{rad}}$  is limited by  $d$ : a longer  $d$  gives more degrees of freedom and increases the maximum possible



(a). Optimized structure for randomly polarized dipole.

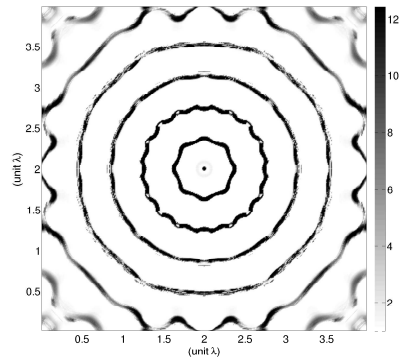


(b).  $H_z$  field of first degenerate mode.

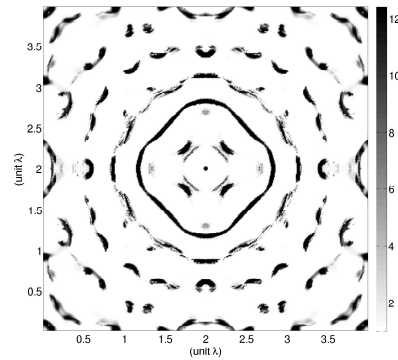


(c).  $H_z$  field of second degenerate mode.

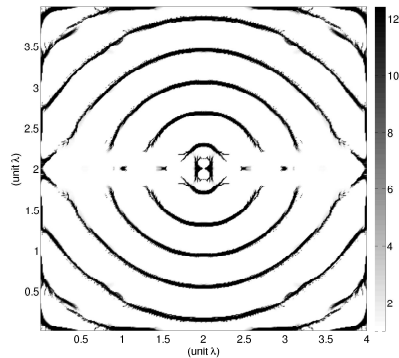
Fig. 7. Optimized doubly-degenerate cavity from Sec. 8.2.3, generated by maximizing the *minimum* LDOS over all in-plane polarizations. Starting from a vacuum initial guess, the optimization discovers a structure with  $C_{3v}$  (3-fold) rotational symmetry (top), which supports doubly-degenerate [74] TE modes whose  $H_z$  fields are plotted at bottom (blue/white/red = negative/zero/positive). (The rectangular FDTD grid slightly breaks the three-fold symmetry and the degeneracy, but the results converge to exact  $C_{3v}$  symmetry and degeneracy as resolution increases.)



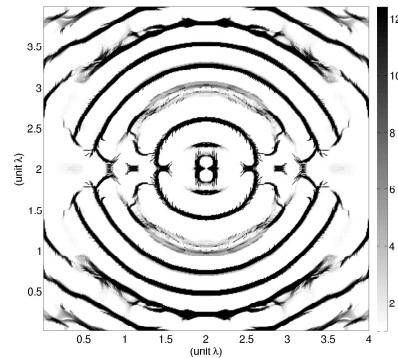
(a). Optimization for TM with  $\tilde{\omega}_2 = 2\tilde{\omega}_1$ .



(b). Optimization for TM with  $\tilde{\omega}_2 = 1.5\tilde{\omega}_1$ .



(c). Optimization for TE with  $\tilde{\omega}_2 = 2\tilde{\omega}_1$ .

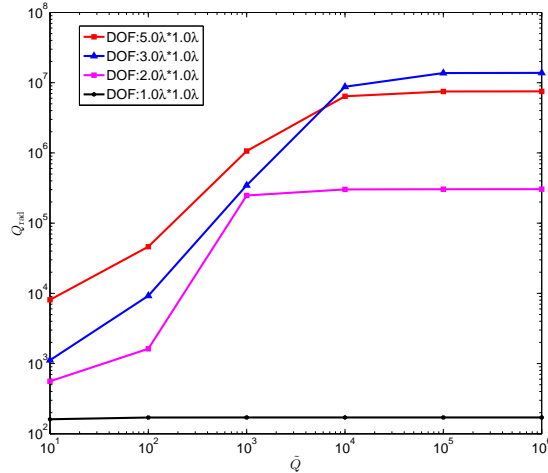


(d). Optimization for TE with  $\tilde{\omega}_2 = 1.5\tilde{\omega}_1$ .

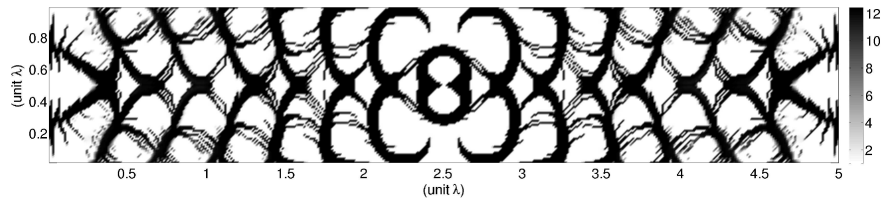
Fig. 8. Two-frequency cavity optimization from Sec. 8.3, for either TM (top) or TE (bottom) polarizations. Left: microcavities which maximize the minimum LDOS at two frequencies  $\tilde{\omega}_1$  and  $\tilde{\omega}_2 = 2\tilde{\omega}_1$ , e.g. for intra-cavity second-harmonic generation applications [79, 81]. Confinement of such integer-multiple frequencies is physically enabled by the fact the concentric Bragg-onion “1d” bandgaps tend to occur at integer-multiple frequencies [14]. A more challenging case is a two-frequency cavity for  $\tilde{\omega}_2 = 1.5\tilde{\omega}_1$ , resulting in the more complicated structures shown at right. (All structures were optimized from vacuum initial guesses.)



$Q_{\text{rad}}$ , as expected. Second, around  $Q_{\text{rad}} \approx 10^7$  the improvement becomes limited by the numerical precision, which prevents the optimization from making progress even though higher  $Q_{\text{rad}}$  should theoretically be possible. For  $d = 5\lambda$ , the resulting structure is shown in Fig. 9(b). [Note that the data in Fig. 9(a) are from the optimization result for the objective  $\varepsilon(\mathbf{x}')\text{LDOS}(\omega, \mathbf{x}')$ .]



(a).  $Q_{\text{rad}}$  vs  $\tilde{Q}$  for 2d thin strip with different length (different DOF).



(b). An optimized structure for 2d thin strip from vacuum initial guess.

Fig. 9. 2D TE optimization for thin strips with fixed width (geometry sketched in Fig. 3b). Fig. (a):  $Q_{\text{rad}}$  vs.  $\tilde{Q}$  for 2d thin strips with same width ( $\lambda$ ) but different length ( $d = \lambda, 2\lambda, 3\lambda, 5\lambda$ ). As  $\tilde{Q}$  is increased in the optimization, higher  $Q_{\text{rad}}$  are obtained until  $Q_{\text{rad}}$  is limited by the degrees of freedom. As the degrees of freedom increase,  $Q_{\text{rad}}$  first gets bigger, but becomes saturated at some level around  $10^7$  due to numerical precision. Fig. (b): An optimized 2d thin-strip structure with width  $\lambda$  and length  $d = 5\lambda$ .

### 8.5. 3D case

In this section, we present the optimization results for the TE-like polarization in the 3d slab setting (Fig. 10). We also briefly discuss some variations on the optimized structure and its comparison with an air-slot cavity.

#### 8.5.1. 3D slab optimization

With the optimization tools developed in the previous sections, we run large-scale simulations on 3d slab case (with an in-plane dipole source, coupling to even-symmetry “TE-like” [14] resonances). Here we choose the dimensions of the slab to be  $3\lambda - 3\lambda - 0.19\lambda$ , where the thickness is  $0.19\lambda = 0.67(\lambda/n)$ . A sketch of the physical model is shown in Fig. 10(a), and the real computation domain (with mirror-symmetry reductions) is illustrated in Fig. 10(b). The optimization

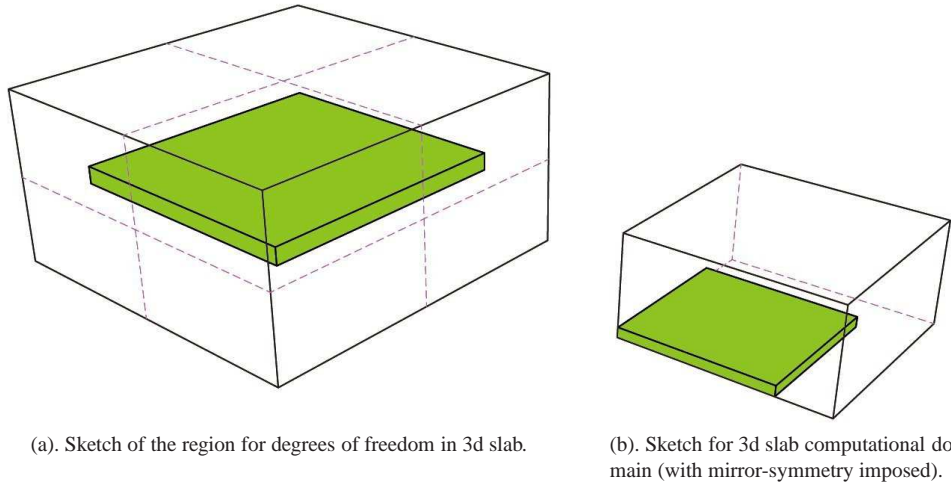


Fig. 10. We wish to optimize a microcavity in an air-membrane Si slab in Sec. 8.5, with the effective computational domain depicted in (a), where the degrees of freedom are every pixel in the 2d pattern of the slab cross-section for a fixed thickness. Since all 2d single-polarization optimizations found structures with two mirror symmetry planes, we can reduce the computational domain to 1/8 the volume (b) by imposing these mirror symmetries along with the vertical mirror symmetry.

discovers a structure (Fig. 11) with quality factor  $Q = 30000$  and extremely small mode volume  $V = 0.06(\lambda/n)^3$ . [This result is obtained from optimizations with absorption  $\tilde{Q}$  gradually increasing from 10 to  $10^4$ . The optimization discovers structures with radiation  $Q = 1.18 \times 10^4$  at  $\tilde{Q} = 100$ , with radiation  $Q = 2.55 \times 10^4$  at  $\tilde{Q} = 1000$ , and with radiation  $Q = 2.98 \times 10^4$  at  $\tilde{Q} = 10^4$ .]

A comparison with other large- or small-scale optimization work, such as 2.5d optimization [72],  $L_3$ -type cavity [20] and  $H_0$ -type cavity [1] optimization are given in table 1. Clearly, the optimization was able to achieve four times smaller mode volume than the smallest mode volume (for  $Q$  within one order of magnitude of ours) we found in the literature [1]. (Note that  $L_3$ -type [20] and  $H_0$ -type [1] cavities are designed from small-scale optimization and can be fabricated; while both 2.5d optimization [72] and our results are purely theoretical and computational investigation from a large-scale optimization perspective.)

Table 1. Comparison of  $Q$  and  $V$  for structures from our result and the literature.

Optimization	Quality Factor $Q$	Mode Volume $V (\lambda/n)^3$
2.5d optimization [72]	8000	0.32
$L_3$ -type cavity optimization [20]	100000	0.70
$H_0$ -type cavity optimization [1]	280000	0.23
LDOS optimization	30000	0.06

### 8.5.2. Sensitivity to small features

In the optimized structure, there are some filament-like structures which would be difficult to fabricate at infrared scales. As a first assessment of the importance of these tiny features, we

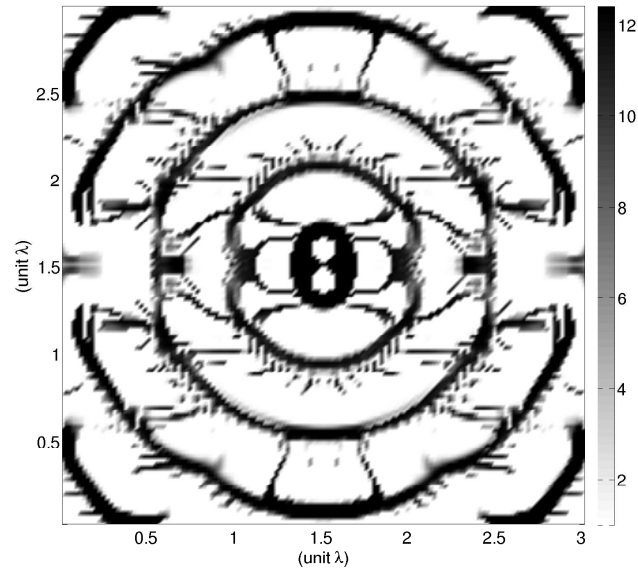


Fig. 11. Optimized pattern for a 3d slab from vacuum initial guess (Sec. 8.5.1) with dimensions  $3\lambda-3\lambda-0.19\lambda$ :  $Q=30000$  and  $V = 0.06(\lambda/n)^3$ .

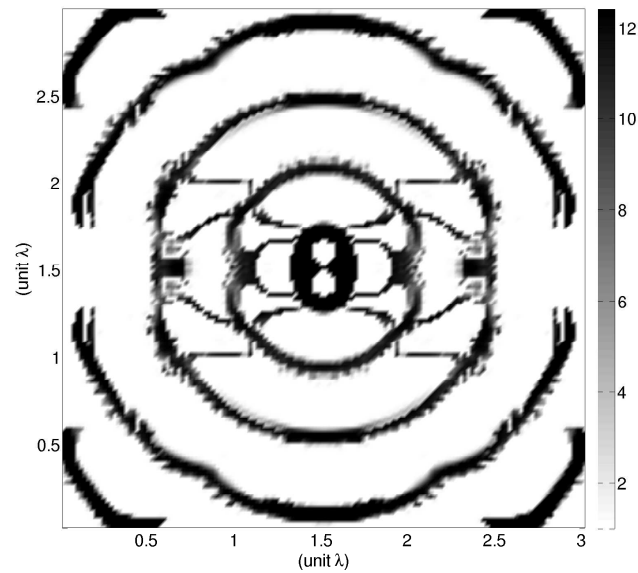


Fig. 12. 3d slab structure after manually removing tiny features in Fig. 11 (Sec. 8.5.2):  $Q = 10000$  and  $V = 0.06(\lambda/n)^3$ .

manually remove the filaments and obtain a structure (Fig. 12) with  $Q = 10000$  and roughly same  $V$ . Instead of post-processing the structure in this way, which has the disadvantage of no longer being a local optimum, future work should consider suppress these tiny features by imposing some explicit constraints during optimization or by some regularization and projection [3] as further discussed in Sec. 9.

### 8.5.3. Comparison with air-slot cavity

All these cavities listed in table 1 are dielectric cavities. In other words, the centers of these cavities are high-dielectric materials (Si and GaAs) and these cavities are useful for dipoles/emitters lying in these materials. It is also reflected in our unit of mode volume. For example, the mode volume of the cavity we obtain is  $V = 0.06(\lambda/n)^3 = 0.06(\lambda/n_{\text{Si}})^3$ .

It is known that air-slot cavities [85–88] can have extremely small volumes. For example, Nomura [86] reported an air-slot cavity with  $Q = 4.8 \times 10^6$  and  $V = 0.015(\lambda/n_{\text{air}})^3$ . Although 0.015 is smaller than 0.06, these two kinds of cavities are not comparable in two ways. First, the two mode volumes are in different units  $(\lambda/n_{\text{air}})^3$  versus  $(\lambda/n_{\text{Si}})^3$ . In our units, their modal volume is  $0.65(\lambda/n_{\text{Si}})^3$ . Second, these two types of cavities are for different applications: air-slot cavities are useful for emitters lying in air, while the semiconductor-based cavities are designed for emitters lying in Si and GaAs.

If the application is for emitters lying in air, in theory, we can also introduce an infinitesimal air-slot at the center, oriented perpendicular to the electric field, into our structure. As discussed in [85], after the introduction of an air-slot, the mode volume decreases by a factor of  $(n_{\text{Si}}/n_{\text{air}})^2$  without changing  $Q$ . In our case, this factor is about 12.4, and the new mode volume is  $4.8 \times 10^{-3}(\lambda/n_{\text{Si}})^3 = 1.1 \times 10^{-4}(\lambda/n_{\text{air}})^3$ . Because the resolution we used (46-pixel per wavelength in air) is not that high, the optimization discovers a dielectric cavity, instead of one with air-slot type. In future work, one could run the optimizations with high resolutions (at least in 2d cases) to investigate whether air-slot structures can be discovered; at such high resolution, the regularized figure of merit of Eq. (21) discussed in Sec. 5.2 becomes essential.

## 9. Manufacturability: Projection and Regularization

The primary purpose of this paper is to improve the mathematical formulation of the microcavity problem to make it amenable to large-scale topology optimization. It is well known, however, that such topology optimization can sometimes lead to non-manufacturable designs, due to three problems: regions of intermediate  $\epsilon$  values, fine features (such as the “hairs” in Fig. 11), and extreme sensitivity to variations in the design parameters. Fortunately, there are a variety of techniques to correct these difficulties, which could easily be combined with our LDOS formulation, and we briefly review these projection, regularization and robustification techniques here for the benefit of future work.

In cases where the optimization does not lead to “bang-bang” solutions, i.e. where there are large regions of intermediate values that do not correspond to available materials, one can use a variety of penalization techniques that add a penalty (increased as needed) to the objective for intermediate  $\epsilon$  values [3, 89]. Another possibility is to use a level-set method [4, 90, 91], which guarantee binary solutions (except possibly on a set of measure zero, since intermediate values are typically used at the level set boundary in order to ensure continuity). A third possibility is the SIMP (Solid Isotropic Material with Penalization) [3], a re-parameterization of the problem that is used in conjunction with the filtering techniques described below in order to project the solution towards a bang-bang result.

To eliminate small features, a common solution is to apply a smoothing filter to the degrees of freedom (combined with a smoothed Heaviside projection to transform back towards

bang-bang designs) [2, 5–8, 92–95]. This is often viewed as a regularization to ensure that the problem is well-posed in the sense that the optimum should converge with resolution (rather than yielding finer and finer features) [3]. Another approach that can eliminate small features as well as designs with extreme parameter sensitivity is “robust” optimization, in which one typically optimizes a worst-case design over a set of parameter uncertainties (e.g. a small “noise” added to each pixel) [67, 96, 97], which has been shown in some cases to also eliminate small features in topology optimization for photonics [68, 98, 99].

## 10. Conclusion

In this paper, we presented a novel formulation for large-scale optimization of optical cavities via frequency-averaged LDOS. With this formulation, we obtained various 2d and 3d cavity-optimization results for different applications. Our results show that several times smaller modal volume is theoretically possible for silicon cavities compared to previous work, without sacrificing  $Q$ . Many other possibilities present themselves for future work. First, the slab thickness in the 3d optimization (Sec. 8.5) is fixed, but one can further extend our approach by allowing the slab thickness to be one additional degree of freedom to improve the figure of merit. Second, although we focused here on silicon-based cavities, one can easily apply our approach to other materials (such as lower-index diamond for visible frequencies [100, 101]). Since silicon cavity design has been so heavily studied and many adequate designs are already known, it is especially in the context of new materials systems that optimization can be valuable. Third, instead of maximizing LDOS, one can maximize the frequency-averaged *total* DOS over the whole computational cell, which is known to be related to the bounds of the light trapping in photovoltaic problems [102, 103]. Fourth, one can extend our work on multi-frequency optimization (Sec. 8.3) by defining a more specialized figure of merit for nonlinear frequency conversion that includes an overlap factor [79, 81] for the two modes.

## Appendix A: Computation of $f'(\omega, \mathbf{x}')$

In this section, we compute  $f'(\omega, \mathbf{x}')$ , the differentiation  $\partial/\partial\omega$  of  $f(\omega, \mathbf{x}')$  from Eq. (12), to show that the residue of Eq. (21) can be evaluated without any additional matrix solves. Differentiating on both sides of Eq. (8), we have

$$\begin{aligned} \mathcal{M}(\varepsilon, \omega) \frac{\partial \mathbf{E}(\mathbf{x}, \omega)}{\partial \omega} + \frac{\partial \mathcal{M}(\varepsilon, \omega)}{\partial \omega} \mathbf{E}(\mathbf{x}, \omega) &= i\mathbf{J}(\mathbf{x}) \\ \implies \frac{\partial \mathbf{E}(\mathbf{x}, \omega)}{\partial \omega} &= \mathcal{M}^{-1} \left( i\mathbf{J}(\mathbf{x}) - \frac{\partial \mathcal{M}(\varepsilon, \omega)}{\partial \omega} \mathbf{E}(\mathbf{x}, \omega) \right) = \mathcal{M}^{-1} (i\mathbf{J}(\mathbf{x}) + 2\omega\varepsilon(\mathbf{x})\mathbf{E}(\mathbf{x}, \omega)). \end{aligned} \quad (28)$$

Now differentiating on Eq. (12), we have

$$\begin{aligned} f'(\omega, \mathbf{x}') &= -\frac{6}{\pi} \int \mathbf{J}^*(\mathbf{x}) \frac{\partial \mathbf{E}(\mathbf{x}, \omega)}{\partial \omega} d\mathbf{x} \\ &= -\frac{6}{\pi} \left[ \int \frac{1}{\omega} \mathbf{J}^*(\mathbf{x}) \mathcal{M}^{-1}(i\omega\mathbf{J}(\mathbf{x})) d\mathbf{x} + 2\omega \int \mathbf{J}^*(\mathbf{x}) \mathcal{M}^{-1} \varepsilon(\mathbf{x}) \mathbf{E}(\mathbf{x}, \omega) d\mathbf{x} \right] \\ &= -\frac{6}{\pi} \int \frac{1}{\omega} \mathbf{J}^*(\mathbf{x}) \mathbf{E}(\mathbf{x}, \omega) d\mathbf{x} + i\frac{12}{\pi} \int (\mathcal{M}^{-1}(i\omega\mathbf{J}(\mathbf{x})))^T \varepsilon(\mathbf{x}) \mathbf{E}(\mathbf{x}, \omega) d\mathbf{x} \\ &= \frac{f(\omega, \mathbf{x}')}{\omega} + i\frac{12}{\pi} \int \varepsilon(\mathbf{x}) \mathbf{E}^T(\mathbf{x}, \omega) \mathbf{E}(\mathbf{x}, \omega) d\mathbf{x}. \end{aligned} \quad (29)$$

Note that we use the properties that  $\mathcal{M}(\varepsilon, \omega)$  is complex symmetric ( $\mathcal{M} = \mathcal{M}^T$ ) for both real and complex  $\omega$  and  $\mathbf{J}(\mathbf{x})$  is real [ $\mathbf{J}^*(\mathbf{x}) = \mathbf{J}^T(\mathbf{x})$ ].

## Appendix B: Computation of the objective and its gradient

In this section, we derive an efficient expression for the objective  $L = \text{Re} [f(\tilde{\omega} + i\tilde{\Gamma}) - i\tilde{\Gamma}f'(\tilde{\omega} + i\tilde{\Gamma})]$  defined in Sec. 5.2 as well as its gradient. The gradient of the objective with respect to the design parameters is calculated with standard adjoint methods [54] in order to minimize the number of required matrix solves.

Let us denote its complex version by  $\mathcal{L} = f(\tilde{\omega} + i\tilde{\Gamma}) - i\tilde{\Gamma}f'(\tilde{\omega} + i\tilde{\Gamma})$ . From Eq. (12) and Eq. (22), we can simplify it as

$$\begin{aligned}\mathcal{L} &= f(\tilde{\omega} + i\tilde{\Gamma}) - i\tilde{\Gamma}f'(\tilde{\omega} + i\tilde{\Gamma}) \\ &= f(\tilde{\omega} + i\tilde{\Gamma}) - i\tilde{\Gamma} \left( \frac{f(\tilde{\omega} + i\tilde{\Gamma}, \mathbf{x}')}{\tilde{\omega} + i\tilde{\Gamma}} + i\frac{12}{\pi} \int \varepsilon(\mathbf{x})\mathbf{E}^T(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x} \right) \\ &= \frac{\tilde{\omega}}{\tilde{\omega} + i\tilde{\Gamma}}f(\tilde{\omega} + i\tilde{\Gamma}) + \frac{12}{\pi}\tilde{\Gamma} \int \varepsilon(\mathbf{x})\mathbf{E}^T(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x}. \\ &= \frac{\tilde{\omega}}{\tilde{\omega} + i\tilde{\Gamma}} \left( -\frac{6}{\pi} \hat{\mathbf{e}}_j^* \cdot \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) + \frac{12}{\pi}\tilde{\Gamma} \int \varepsilon(\mathbf{x})\mathbf{E}^T(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x} \right).\end{aligned}\quad (30)$$

In the rest of this section, we compute the gradient of  $\mathcal{L}$  with respect to the design parameter  $\varepsilon_k$ , which is the dielectric constant at  $\mathbf{x} = \mathbf{x}_k$ . To obtain the sensitivity of  $\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})$  to  $\varepsilon_k$ , we differentiate Eq. (17) with respect to  $\varepsilon_k$

$$\begin{aligned}\tilde{\mathcal{M}}(\varepsilon, \tilde{\omega}) \frac{\partial \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})}{\partial \varepsilon_k} + \frac{\partial \tilde{\mathcal{M}}(\varepsilon, \tilde{\omega})}{\partial \varepsilon_k} \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) &= 0 \\ \implies \frac{\partial \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})}{\partial \varepsilon_k} &= \tilde{\mathcal{M}}^{-1} \left[ \tilde{\omega}^2 \left( 1 + \frac{i}{2\tilde{Q}} \right) \delta(\mathbf{x} - \mathbf{x}_k) \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) \right].\end{aligned}\quad (31)$$

Therefore, from Eq. (12), Eq. (22) and Eq. (31), we have

$$\begin{aligned}\frac{\partial \int \mathbf{J}^*(\mathbf{x})\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x}}{\partial \varepsilon_k} &= \int \mathbf{J}^*(\mathbf{x}) \frac{\partial \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})}{\partial \varepsilon_k} d\mathbf{x} \\ &= \tilde{\omega}^2 \left( 1 + \frac{i}{2\tilde{Q}} \right) \int (\tilde{\mathcal{M}}^{-1} \mathbf{J}(\mathbf{x}))^T \delta(\mathbf{x} - \mathbf{x}_k) \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x} \\ &= -i\tilde{\omega} \left( 1 + \frac{i}{2\tilde{Q}} \right) \int \mathbf{E}^T(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) \delta(\mathbf{x} - \mathbf{x}_k) \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x} \\ &= -i\tilde{\omega} \left( 1 + \frac{i}{2\tilde{Q}} \right) \mathbf{E}^T(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \mathbf{E}(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}),\end{aligned}\quad (32)$$

and

$$\begin{aligned}\frac{\partial \int \varepsilon(\mathbf{x})\mathbf{E}^T \mathbf{E} d\mathbf{x}}{\partial \varepsilon_k} &= 2 \int \varepsilon(\mathbf{x})\mathbf{E}^T \frac{\partial \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})}{\partial \varepsilon_k} d\mathbf{x} + \int \delta(\mathbf{x} - \mathbf{x}_k) \mathbf{E}^T(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x} \\ &= 2 \int \varepsilon(\mathbf{x})\mathbf{E}^T \tilde{\mathcal{M}}^{-1} \left[ \tilde{\omega}^2 \left( 1 + \frac{i}{2\tilde{Q}} \right) \delta(\mathbf{x} - \mathbf{x}_k) \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) \right] \\ &\quad + \mathbf{E}^T(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \mathbf{E}(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \\ &= 2\tilde{\omega}^2 \left( 1 + \frac{i}{2\tilde{Q}} \right) \int (\tilde{\mathcal{M}}^{-1} [\varepsilon(\mathbf{x})\mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})])^T \delta(\mathbf{x} - \mathbf{x}_k) \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})d\mathbf{x} \\ &\quad + \mathbf{E}^T(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \mathbf{E}(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}).\end{aligned}\quad (33)$$



Combining Eq. (30), Eq. (32) and Eq. (33), we have

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \varepsilon_k} = & \left( i + \frac{1}{\tilde{Q}} \right) \frac{6}{\pi} \tilde{\omega} \mathbf{E}^T(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \mathbf{E}(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \\ & + \frac{12}{\pi} \frac{\tilde{\omega}^3}{\tilde{Q}} \left( 1 + \frac{i}{2\tilde{Q}} \right) \mathbf{A}^T(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}) \mathbf{E}(\mathbf{x}_k, \tilde{\omega} + i\tilde{\Gamma}), \end{aligned} \quad (34)$$

where  $\mathbf{A}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma})$  satisfies the scattering equation

$$\tilde{\mathcal{M}}(\varepsilon, \tilde{\omega}) \mathbf{A}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}) = \varepsilon(\mathbf{x}) \mathbf{E}(\mathbf{x}, \tilde{\omega} + i\tilde{\Gamma}). \quad (35)$$

Therefore, as usual for adjoint-methods [54], the gradient of  $\mathcal{L}$  with respect to design parameters  $\varepsilon$  can be evaluated with only a single additional matrix solve Eq. (35).

### Acknowledgments

X.L. thanks J. Joannopoulos and B. Zhen for helpful discussions. This work was supported in part by the AFOSR MURI for Complex and Robust On-chip Nanophotonics (Dr. Gernot Pomrenke) grant number FA9550-09-1-0704 and by the Institute for Soldier Nanotechnologies under Contract W911NF-07-D0004. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575.