

## MIT Open Access Articles

### *Fundamental limits to optical response in absorptive systems*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Miller, Owen D., Athanasios G. Polimeridis, M. T. Homer Reid, Chia Wei Hsu, Brendan G. DeLacy, John D. Joannopoulos, Marin Soljacic, and Steven G. Johnson. "Fundamental Limits to Optical Response in Absorptive Systems." *Optics Express* 24, no. 4 (February 9, 2016): 3329.

**As Published:** <http://dx.doi.org/10.1364/OE.24.003329>

**Publisher:** Optical Society of America

**Persistent URL:** <http://hdl.handle.net/1721.1/101261>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of use:** Creative Commons Attribution-Noncommercial-Share Alike



# Fundamental limits to optical response in absorptive systems

Owen D. Miller,<sup>1,\*</sup> Athanasios G. Polimeridis,<sup>2</sup> M. T. Homer Reid,<sup>1</sup>  
Chia Wei Hsu,<sup>3</sup> Brendan G. DeLacy,<sup>4</sup> John D. Joannopoulos,<sup>5</sup> Marin  
Soljačić,<sup>5</sup> and Steven G. Johnson<sup>1</sup>

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

<sup>2</sup>*Skolkovo Institute of Science and Technology, 143025 Moscow Region, Russia*

<sup>3</sup>*Department of Applied Physics, Yale University, New Haven, CT, 06520, USA*

<sup>4</sup>*U.S. Army Edgewood Chemical Biological Center, Research and Technology Directorate, Aberdeen Proving Ground, MD, 21010, USA*

<sup>5</sup>*Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA*

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

**Abstract:** At visible and infrared frequencies, metals show tantalizing promise for strong subwavelength resonances, but material loss typically dampens the response. We derive fundamental limits to the optical response of absorptive systems, bounding the largest enhancements possible given intrinsic material losses. Through basic conservation-of-energy principles, we derive geometry-independent limits to per-volume absorption and scattering rates, and to local-density-of-states enhancements that represent the power radiated or expended by a dipole near a material body. We provide examples of structures that approach our absorption and scattering limits at any frequency; by contrast, we find that common “antenna” structures fall far short of our radiative LDOS bounds, suggesting the possibility for significant further improvement. Underlying the limits is a simple metric,  $|\chi|^2/\text{Im}\chi$  for a material with susceptibility  $\chi$ , that enables broad technological evaluation of lossy materials across optical frequencies.

© 2016 Optical Society of America

**OCIS codes:** (260.3910) Metal optics; (250.5403) Plasmonics; (310.6805) Theory and design.

---

## References and links

1. E. Ozbay, “Plasmonics: merging photonics and electronics at nanoscale dimensions,” *Science* **311**, 189–193 (2006).
2. S. A. Maier, *Plasmonics: Fundamentals and Applications* (Springer Science & Business Media, 2007).
3. A. Boltasseva and H. A. Atwater, “Low-loss plasmonic metamaterials,” *Science* **331**, 290–291 (2011).
4. H. A. Atwater and A. Polman, “Plasmonics for improved photovoltaic devices,” *Nat. Mater.* **9**, 205–213 (2010).
5. G. V. Naik, J. Kim, and A. Boltasseva, “Oxides and nitrides as alternative plasmonic materials in the optical range [Invited],” *Opt. Mater. Express* **1**, 1090–1099 (2011).
6. P. Tassin, T. Koschny, M. Kafesaki, and C. M. Soukoulis, “A comparison of graphene, superconductors and metals as conductors for metamaterials and plasmonics,” *Nat. Photonics* **6**, 259–264 (2012).
7. M. D. Arnold and M. G. Blaber, “Optical performance and metallic absorption in nanoplasmonic systems,” *Opt. Express* **17**, 3835–3847 (2009).
8. J. B. Khurgin and G. Sun, “In search of the elusive lossless metal,” *Appl. Phys. Lett.* **96**, 181102 (2010).
9. A. Raman, W. Shin, and S. Fan, “Upper bound on the modal material loss rate in plasmonic and metamaterial systems,” *Phys. Rev. Lett.* **110**, 183901 (2013).
10. J. B. Khurgin, “How to deal with the loss in plasmonics and metamaterials,” *Nat. Nanotechnol.* **10**, 2–6 (2015).
11. L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University, 2012), 2nd ed.

12. K. Joulain, R. Carminati, J.-P. Mulet, and J.-J. Greffet, "Definition and measurement of the local density of electromagnetic states close to an interface," *Phys. Rev. B* **68**, 245405 (2003).
13. O. Martin and N. Piller, "Electromagnetic scattering in polarizable backgrounds," *Phys. Rev. E* **58**, 3909–3915 (1998).
14. G. D'Aguanno, 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).
15. A. Oskooi and S. G. Johnson, "Electromagnetic wave source conditions," in "Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology," A. Taflove, A. Oskooi, and S. G. Johnson, eds. (Artech, 2013), chap. 4, pp. 65–100.
16. X. Liang and S. G. Johnson, "Formulation for scalable optimization of microcavities via the frequency-averaged local density of states," *Opt. Express* **21**, 30812–30841 (2013).
17. E. M. Purcell, "On the absorption and emission of light by interstellar grains," *Astrophys. J.* **158**, 433–440 (1969).
18. C. Sohl, M. Gustafsson, and G. Kristensson, "Physical limitations on metamaterials: restrictions on scattering and absorption over a frequency interval," *J. Phys. D: Appl. Phys.* **40**, 7146–7151 (2007).
19. C. Sohl, M. Gustafsson, and G. Kristensson, "Physical limitations on broadband scattering by heterogeneous obstacles," *J. Phys. A Math. Theor.* **40**, 11165–11182 (2007).
20. M. Gustafsson, C. Sohl, and G. Kristensson, "Physical limitations on antennas of arbitrary shape," *Proc. R. Soc. A Math. Phys. Eng. Sci.* **463**, 2589–2607 (2007).
21. D.-H. Kwon and D. M. Pozar, "Optimal characteristics of an arbitrary receive antenna," *IEEE Trans. Antennas Propag.* **57**, 3720–3727 (2009).
22. I. Liberal, Y. Ra'di, R. Gonzalo, I. Ederra, S. A. Tretyakov, and R. W. Ziolkowski, "Least upper bounds of the powers extracted and scattered by bi-anisotropic particles," *IEEE Trans. Antennas Propag.* **62**, 4726–4735 (2014).
23. H. A. Wheeler, "Fundamental limitations of small antennas," *Proc. IRE* **35**, 1479–1484 (1947).
24. L. J. Chu, "Physical limitations of omni-directional antennas," *J. Appl. Phys.* **19**, 1163–1175 (1948).
25. J. S. McLean, "A re-examination of the fundamental limits on the radiation Q of electrically small antennas," *IEEE Trans. Antennas Propag.* **44**, 672–676 (1996).
26. F. Wang and Y. R. Shen, "General properties of local plasmons in metal nanostructures," *Phys. Rev. Lett.* **97**, 206806 (2006).
27. J.-P. Hugonin, M. Besbes, and P. Ben-Abdallah, "Fundamental limits for light absorption and scattering induced by cooperative electromagnetic interactions," *Phys. Rev. B* **91**, 180202 (2015).
28. P. R. West, S. Ishii, G. V. Naik, N. K. Emani, V. M. Shalae, and A. Boltasseva, "Searching for better plasmonic materials," *Laser Photon. Rev.* **4**, 795–808 (2010).
29. G. V. Naik, V. M. Shalae, and A. Boltasseva, "Alternative plasmonic materials: beyond gold and silver," *Adv. Mater.* **25**, 3264–3294 (2013).
30. S. Law, L. Yu, A. Rosenberg, and D. Wasserman, "All-semiconductor plasmonic nanoantennas for infrared sensing," *Nano Lett.* **13**, 4569–4574 (2013).
31. R. D. Averitt, S. L. Westcott, and N. J. Halas, "Linear optical properties of gold nanoshells," *J. Opt. Soc. Am. B* **16**, 1824–1832 (1999).
32. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (John Wiley & Sons, 1983).
33. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer, 1988).
34. D. Pines, *Elementary Excitations in Solids: Lectures on Phonons, Electrons, and Plasmons*, vol. 5 (W. A. Benjamin, 1964).
35. V. J. Keast, "Ab initio calculations of plasmons and interband transitions in the low-loss electron energy-loss spectrum," *J. Electron Spectros. Relat. Phenomena* **143**, 97–104 (2005).
36. D. R. Lytle, P. S. Carney, J. C. Schotland, and E. Wolf, "Generalized optical theorem for reflection, transmission, and extinction of power for electromagnetic fields," *Phys. Rev. E* **71**, 056610 (2005).
37. R. G. Newton, "Optical theorem and beyond," *Am. J. Phys.* **44**, 639–642 (1976).
38. J. D. Jackson, *Classical Electrodynamics, 3rd Ed.* (John Wiley & Sons, 1999).
39. J. A. Scholl, A. L. Koh, and J. A. Dionne, "Quantum plasmon resonances of individual metallic nanoparticles," *Nature* **483**, 421–427 (2012).
40. C. Ciraci, R. T. Hill, J. J. Mock, Y. Urzhumov, A. I. Fernández-Domínguez, S. A. Maier, J. B. Pendry, A. Chilkoti, and D. R. Smith, "Probing the ultimate limits of plasmonic enhancement," *Science* **337**, 1072–1074 (2012).
41. M. S. Eggleston, K. Messer, L. Zhang, E. Yablonovitch, and M. C. Wu, "Optical antenna enhanced spontaneous emission," *Proc. Natl. Acad. Sci.* **112**, 1704–1709 (2015).
42. R. C. Hansen, *Electrically Small, Superdirective, and Superconducting Antennas* (John Wiley & Sons, 2006).
43. R. E. Hamam, A. Karalis, J. D. Joannopoulos, and M. Soljačić, "Coupled-mode theory for general free-space resonant scattering of waves," *Phys. Rev. A* **75**, 053801 (2007).
44. Z. Ruan and S. Fan, "Design of subwavelength superscattering nanospheres," *Appl. Phys. Lett.* **98**, 043101 (2011).
45. O. D. Miller, C. W. Hsu, M. T. H. Reid, W. Qiu, B. G. DeLacy, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, "Fundamental limits to extinction by metallic nanoparticles," *Phys. Rev. Lett.* **112**, 123903 (2014).

46. R. Fuchs, "Theory of the optical properties of ionic crystal cubes," *Phys. Rev. B* **11**, 1732–1740 (1975).
47. R. Fuchs and S. H. Liu, "Sum rule for the polarizability of small particles," *Phys. Rev. B* **14**, 5521–5522 (1976).
48. G. W. Milton, personal communication.
49. G. W. Milton, "Bounds on the complex permittivity of a two-component composite material," *J. Appl. Phys.* **52**, 5286–5293 (1981).
50. D. J. Bergman, "Bounds for the complex dielectric constant of a two-component composite material," *Phys. Rev. B* **23**, 3058–3065 (1981).
51. W. C. Chew, *Waves and Fields in Inhomogeneous Media* (IEEE, 1995).
52. A. Welters, Y. Avniel, and S. G. Johnson, "Speed-of-light limitations in passive linear media," *Phys. Rev. A* **90**, 023847 (2014).
53. A. G. Polimeridis, M. T. H. Reid, S. G. Johnson, J. K. White, and A. W. Rodriguez, "On the computation of power in volume integral equation formulations," *IEEE Trans. Antennas Propag.* **63**, 611–620 (2015).
54. J. A. Kong, "Theorems of bianisotropic media," *Proc. IEEE* **60**, 1036–1046 (1972).
55. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference, and Diffraction of Light* (Cambridge University, 1999), 7th ed.
56. H. Hashemi, C.-W. Qiu, A. P. McCauley, J. D. Joannopoulos, and S. G. Johnson, "Diameter-bandwidth product limitation of isolated-object cloaking," *Phys. Rev. A* **86**, 013804 (2012).
57. S. Nie and S. R. Emory, "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering," *Science* **275**, 1102–1106 (1997).
58. K. Kneipp, Y. Wang, H. Kneipp, L. Perelman, I. Itzkan, R. Dasari, and M. Feld, "Single molecule detection using surface-enhanced Raman scattering (SERS)," *Phys. Rev. Lett.* **78**, 1667–1670 (1997).
59. T. S. van Zanten, A. Cambi, M. Koopman, B. Joosten, C. G. Figdor, and M. F. Garcia-Parajo, "Hotspots of gpi-anchored proteins and integrin nanoclusters function as nucleation sites for cell adhesion," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 18557–18562 (2009).
60. L. Schermelleh, R. Heintzmann, and H. Leonhardt, "A guide to super-resolution fluorescence microscopy," *J. Cell Biol.* **190**, 165–175 (2010).
61. M. Laroche, R. Carminati, and J.-J. Greffet, "Near-field thermophotovoltaic energy conversion," *J. Appl. Phys.* **100**, 063704 (2006).
62. S. Basu and M. Francoeur, "Maximum near-field radiative heat transfer between thin films," *Appl. Phys. Lett.* **98**, 243120 (2011).
63. E. K. Lau, A. Lakhani, R. S. Tucker, and M. C. Wu, "Enhanced modulation bandwidth of nanocavity light emitting devices," *Opt. Express* **17**, 7790–7799 (2009).
64. F. Wijnands, J. B. Pendry, F. J. Garcia-Vidal, P. M. Bell, P. J. Roberts, and L. M. Moreno, "Green's functions for Maxwell's equations: application to spontaneous emission," *Opt. Quantum Electron.* **29**, 199–216 (1997).
65. Y. Xu, R. K. Lee, and A. Yariv, "Quantum analysis and the classical analysis of spontaneous emission in a microcavity," *Phys. Rev. A* **61**, 033807 (2000).
66. L. Novotny, "Effective wavelength scaling for optical antennas," *Phys. Rev. Lett.* **98**, 266802 (2007).
67. C.-T. Tai, *Dyadic Green Functions in Electromagnetic Theory* (IEEE, 1994), 2nd ed.
68. E. J. Rothwell and M. J. Cloud, *Electromagnetics* (CRC, 2001).
69. K. Joulain, J.-P. Mulet, F. Marquier, R. Carminati, and J.-J. Greffet, "Surface electromagnetic waves thermally excited: radiative heat transfer, coherence properties and Casimir forces revisited in the near field," *Surf. Sci. Rep.* **57**, 59–112 (2005).
70. I. M. Gelfand and S. V. Fomin, *Calculus of Variations* (Dover, 2000).
71. K. Kreutz-Delgado, "The complex gradient operator and the CR-calculus," arXiv:0906.4835 (2009).
72. R. Remmert, *Theory of Complex Functions* (Springer-Verlag, 1991).
73. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University, 2004).
74. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, 2006), 2nd ed.
75. J.-B. Hiriart-Urruty, "Global optimality conditions in maximizing a convex quadratic function under convex quadratic constraints," *J. Glob. Optim.* **21**, 443–453 (2001).
76. L. N. Trefethen and D. Bau, *Numerical Linear Algebra* (Society for Industrial and Applied Mathematics, 1997).
77. F. Capolino, ed., *Metamaterials Handbook* (CRC, 2009).
78. W. G. Spitzer, D. Kleinman, and D. Walsh, "Infrared properties of hexagonal silicon carbide," *Phys. Rev.* **113**, 127–132 (1959).
79. E. D. Palik, *Handbook of Optical Constants of Solids* (Elsevier Science, 1998).
80. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT, Cambridge, MA, 2009), 3rd ed.
81. A. P. Hibbins, J. R. Sambles, and C. R. Lawrence, "Surface plasmon-polariton study of the optical dielectric function of titanium nitride," *J. Mod. Opt.* **45**, 2051–2062 (1998).
82. S. Franzen, "Surface plasmon polaritons and screened plasma absorption in indium tin oxide compared to silver and gold," *J. Phys. Chem. C* **112**, 6027–6032 (2008).
83. E. Sachet, C. T. Shelton, J. S. Harris, B. E. Gaddy, D. L. Irving, S. Curtarolo, B. F. Donovan, P. E. Hopkins, P. A. Sharma, A. L. Sharma, J. Ihlefeld, S. Franzen, and J.-P. Maria, "Dysprosium-doped cadmium oxide as a gateway

- material for mid-infrared plasmonics,” *Nat. Mater.* **14**, 414–420 (2015).
84. B. D. Busbee, S. O. Obare, and C. J. Murphy, “An improved synthesis of high-aspect-ratio gold nanorods,” *Adv. Mater.* **15**, 414–416 (2003).
  85. P. B. Johnson and R. W. Christy, “Optical constants of the noble metals,” *Phys. Rev. B* **6**, 4370–4379 (1972).
  86. Y. Wu, C. Zhang, N. M. Estakhri, Y. Zhao, J. Kim, M. Zhang, X. X. Liu, G. K. Pribil, A. Alù, C. K. Shih, and X. Li, “Intrinsic optical properties and enhanced plasmonic response of epitaxial silver,” *Adv. Mater.* **26**, 6106–6110 (2014).
  87. S. Vassant, A. Archambault, F. Marquier, F. Pardo, U. Gennser, A. Cavanna, J. L. Pelouard, and J. J. Greffet, “Epsilon-near-zero mode for active optoelectronic devices,” *Phys. Rev. Lett.* **109**, 237401 (2012).
  88. S. Campione, I. Brener, and F. Marquier, “Theory of epsilon-near-zero modes in ultrathin films,” *Phys. Rev. B* **91**, 121408 (2015).
  89. M. I. Tribelsky and B. S. Lukyanchuk, “Anomalous light scattering by small particles,” *Phys. Rev. Lett.* **97**, 263902 (2006).
  90. M. I. Tribelsky, “Anomalous light absorption by small particles,” *Europhys. Lett.* **94**, 14004 (2011).
  91. 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).
  92. V. V. Klimov, M. Ducloy, and V. S. Letokhov, “Spontaneous emission of an atom in the presence of nanobodies,” *Quantum Electron.* **31**, 569–586 (2001).
  93. L. C. Davis, “Electrostatic edge modes of a dielectric wedge,” *Phys. Rev. B* **14**, 5523–5525 (1976).
  94. J. Andersen and V. Solodukhov, “Field behavior near a dielectric wedge,” *IEEE Trans. Antennas Propag.* **26**, 598–602 (1978).
  95. O. D. Miller, S. G. Johnson, and A. W. Rodriguez, “Effectiveness of thin films in lieu of hyperbolic metamaterials in the near field,” *Phys. Rev. Lett.* **112**, 157402 (2014).
  96. M. T. H. Reid and S. G. Johnson, “Efficient computation of power, force, and torque in BEM scattering calculations,” *IEEE Trans. Antennas Propag.* **63**, 3588–3598 (2015).
  97. M. T. H. Reid, “Scuff-em: free, open-source boundary-element software,” <http://homerreid.com/scuff-EM>.
  98. R. F. Harrington, *Field Computation by Moment Methods* (IEEE, 1993).
  99. J. E. Sipe, “New Green-function formalism for surface optics,” *J. Opt. Soc. Am. B.* **4**, 481–489 (1987).
  100. E. J. Zeman and G. C. Schatz, “An accurate electromagnetic theory study of surface enhancement factors for silver, gold, copper, lithium, sodium, aluminum, gallium, indium, zinc, and cadmium,” *J. Phys. Chem.* **91**, 634–643 (1987).
  101. K. L. Kelly, E. Coronado, L. L. Zhao, and G. C. Schatz, “The optical properties of metal nanoparticles: the influence of size, shape, and dielectric environment,” *J. Phys. Chem. B* **107**, 668–677 (2003).
  102. A. Moroz, “Depolarization field of spheroidal particles,” *J. Opt. Soc. Am. B* **26**, 517–527 (2009).
  103. S. Law, L. Yu, and D. Wasserman, “Epitaxial growth of engineered metals for mid-infrared plasmonics,” *J. Vac. Sci. Technol. B* **31**, 03C121 (2013).
  104. S. G. Johnson, “The nlopt nonlinear-optimization package,” <http://ab-initio.mit.edu/nlopt>.
  105. P. Kaelo and M. M. Ali, “Some variants of the controlled random search algorithm for global optimization,” *J. Optim. Theory Appl.* **130**, 253–264 (2006).
  106. W. L. Price, “Global optimization by controlled random search,” *J. Optim. Theory Appl.* **40**, 333–348 (1983).
  107. E. Anquillare, O. D. Miller, C. W. Hsu, B. G. DeLacy, J. D. Joannopoulos, S. G. Johnson, and M. Soljačić, “Efficient broad- and tunable-bandwidth optical extinction via aspect-ratio-tailored silver nanodisks,” *Submit. arXiv 1510.01768* (2015).
  108. C. David and F. J. Garcia de Abajo, “Surface plasmon dependence on the electron density profile at metal surfaces,” *ACS Nano* **8**, 9558–9566 (2014).
  109. A. M. Gobin, M. H. Lee, N. J. Halas, W. D. James, R. A. Drezek, and J. L. West, “Near-infrared resonant nanoshells for combined optical imaging and photothermal cancer therapy,” *Nano Lett.* **7**, 1929–1934 (2007).
  110. E. C. Dreaden, A. M. Alkilany, X. Huang, C. J. Murphy, and M. A. El-Sayed, “The golden age: gold nanoparticles for biomedicine,” *Chem. Soc. Rev.* **41**, 2740–2779 (2012).
  111. D. Polder and M. Van Hove, “Theory of radiative heat transfer between closely spaced bodies,” *Phys. Rev. B* **4**, 3303–3314 (1971).
  112. S. M. Rytov, Y. A. Kravtsov, and V. I. Tatarskii, *Principles of Statistical Radiophysics* (Springer-Verlag, 1988).
  113. J. B. Pendry, “Radiative exchange of heat between nanostructures,” *J. Phys. Condens. Matter* **11**, 6621–6633 (1999).
  114. R. S. Ottens, V. Quetschke, S. Wise, A. A. Alemi, R. Lundock, G. Mueller, D. H. Reitze, D. B. Tanner, and B. F. Whiting, “Near-field radiative heat transfer between macroscopic planar surfaces,” *Phys. Rev. Lett.* **107**, 014301 (2011).
  115. M. Francoeur, M. P. Menguc, and R. Vaillon, “Near-field radiative heat transfer enhancement via surface phonon polaritons coupling in thin films,” *Appl. Phys. Lett.* **93**, 043109 (2008).
  116. W. Nomura, M. Ohtsu, and T. Yatsui, “Nanodot coupler with a surface plasmon polariton condenser for optical far/near-field conversion,” *Appl. Phys. Lett.* **86**, 181108 (2005).

117. F. Lopez-Tejeira, S. G. Rodrigo, L. Martin-Moreno, F. J. Garcia-Vidal, E. Devaux, T. W. Ebbesen, J. R. Krenn, P. Radko, S. I. Bozhevolnyi, M. U. Gonzalez, J. C. Weeber, and A. Dereux, "Efficient unidirectional nanoslit couplers for surface plasmons," *Nat. Phys.* **3**, 324–328 (2007).
118. A. Baron, E. Devaux, J. C. Rodier, J. P. Hugonin, E. Rousseau, C. Genet, T. W. Ebbesen, and P. Lalanne, "Compact antenna for efficient and unidirectional launching and decoupling of surface plasmons," *Nano Lett.* **11**, 4207–4212 (2011).
119. E. N. Economou, "Surface plasmons in thin films," *Phys. Rev.* **182**, 539–554 (1969).
120. H. N. S. Krishnamoorthy, Z. Jacob, E. Narimanov, I. Kretzschmar, and V. M. Menon, "Topological transitions in metamaterials," *Science* **336**, 205–209 (2012).
121. C. L. Cortes, W. Newman, S. Molesky, and Z. Jacob, "Quantum nanophotonics using hyperbolic metamaterials," *J. Opt.* **14**, 063001 (2012).
122. A. Poddubny, I. Iorsh, P. Belov, and Y. Kivshar, "Hyperbolic metamaterials," *Nat. Photonics* **7**, 948–957 (2013).
123. M. J. D. Powell, "A direct search optimization method that models the objective and constraint functions by linear interpolation," in "Advances in Optimization and Numerical Analysis," (Springer, 1994), pp. 51–67.
124. M. Silveirinha and C. Fernandes, "A new acceleration technique with exponential convergence rate to evaluate periodic Green functions," *IEEE Trans. Antennas Propag.* **53**, 347–355 (2005).
125. F. Capolino, D. Wilton, and W. Johnson, "Efficient computation of the 3d Greens function for the helmholtz operator for a linear array of point sources using the ewald method," *J. Comput. Phys.* **223**, 250–261 (2007).
126. R. A. Shore and A. D. Yaghjian, "Traveling waves on two- and three-dimensional periodic arrays of lossless scatterers," *Radio Sci.* **42**, RS6S21 (2007).
127. X. Liu, T. Starr, A. F. Starr, and W. J. Padilla, "Infrared spatial and frequency selective metamaterial with near-unity absorbance," *Phys. Rev. Lett.* **104**, 207403 (2010).
128. A. Tittl, P. Mai, R. Taubert, D. Dregely, N. Liu, and H. Giessen, "Palladium-based plasmonic perfect absorber in the visible wavelength range and its application to hydrogen sensing," *Nano Lett.* **11**, 4366–4369 (2011).
129. N. Landy, S. Sajuyigbe, J. Mock, D. Smith, and W. Padilla, "Perfect metamaterial absorber," *Phys. Rev. Lett.* **100**, 207402 (2008).
130. R. A. Pala, J. White, E. Barnard, J. Liu, and M. L. Brongersma, "Design of plasmonic thin-film solar cells with broadband absorption enhancements," *Adv. Mater.* **21**, 3504–3509 (2009).
131. K. Aydin, V. E. Ferry, R. M. Briggs, and H. A. Atwater, "Broadband polarization-independent resonant light absorption using ultrathin plasmonic super absorbers," *Nat. Commun.* **2**, 517 (2011).
132. C. W. Hsu, B. Zhen, W. Qiu, O. Shapira, B. G. DeLacy, J. D. Joannopoulos, and M. Soljačić, "Transparent displays enabled by resonant nanoparticle scattering," *Nat. Commun.* **5**, 3152 (2014).
133. D. Peer, J. M. Karp, S. Hong, O. C. Farokhzad, R. Margalit, and R. Langer, "Nanocarriers as an emerging platform for cancer therapy," *Nat. Nanotechnol.* **2**, 751–760 (2007).
134. E. Boisselier and D. Astruc, "Gold nanoparticles in nanomedicine: preparations, imaging, diagnostics, therapies and toxicity," *Chem. Soc. Rev.* **38**, 1759–1782 (2009).
135. W. Qiu, B. G. Delacy, S. G. Johnson, J. D. Joannopoulos, and M. Soljačić, "Optimization of broadband optical response of multilayer nanospheres," *Opt. Express* **20**, 18494–18504 (2012).
136. M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, 1980).
137. G. K. Pedersen, *Analysis Now* (Springer, 1989).
138. T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, 1980), 2nd ed.
139. W. D. Heiss, "Exceptional points of non-Hermitian operators," *J. Phys. A: Math. Gen.* **37**, 2455–2464 (2004).
140. L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: the Behavior of Nonnormal Matrices and Operators* (Princeton University, 2005).
141. C. W. Hsu, B. G. DeLacy, S. G. Johnson, J. D. Joannopoulos, and M. Soljačić, "Theoretical criteria for scattering dark states in nanostructured particles," *Nano Lett.* **14**, 2783–2788 (2014).
142. U. Fano, "Effects of configuration interaction on intensities and phase shifts," *Phys. Rev.* **124**, 1866–1878 (1961).
143. K. Petermann, "Calculated spontaneous emission factor for double-heterostructure injection lasers with gain-induced waveguiding," *IEEE J. Quantum Electron.* **15**, 566–570 (1979).
144. A. Siegman, "Excess spontaneous emission in non-Hermitian optical systems. i. laser amplifiers," *Phys. Rev. A* **39**, 1253–1263 (1989).
145. O. D. Miller, S. G. Johnson, and A. W. Rodriguez, "Shape-independent limits to near-field radiative heat transfer," *Phys. Rev. Lett.* **115**, 204302 (2015).
146. J. Rahola, "On the eigenvalues of the volume integral operator of electromagnetic scattering," *SIAM J. Sci. Comput.* **21**, 1740–1754 (2000).
147. S. Esterhazy, D. Liu, A. Cerjan, L. Ge, K. G. Makris, A. D. Stone, J. M. Melenk, S. G. Johnson, and S. Rotter, "Scalable numerical approach for the steady-state ab initio laser theory," *Phys. Rev. A* **90**, 023816 (2014).
148. F. Ouyang and M. Isaacson, "Surface plasmon excitation of objects with arbitrary shape and dielectric constant," *Philos. Mag. B* **60**, 481–492 (1989).
149. I. D. Mayergoyz, D. R. Fredkin, and Z. Zhang, "Electrostatic (plasmon) resonances in nanoparticles," *Phys. Rev. B* **72**, 155412 (2005).
150. P. D. Lax and R. S. Phillips, *Scattering Theory* (Academic, 1990).

151. M. Zworski, “Resonances in physics and geometry,” *Not. AMS* **46**, 319–328 (1999).
152. A. G. Polimeridis, J. F. Villena, L. Daniel, and J. K. White, “Stable FFT-JVIE solvers for fast analysis of highly inhomogeneous dielectric objects,” *J. Comput. Phys.* **269**, 280–296 (2014).
153. L. E. Sun and W. C. Chew, “A novel formulation of the volume integral equation for electromagnetic scattering,” *Waves in Random and Complex Media* **19**, 162–180 (2009).
154. A. B. Samokhin, *Integral Equations and Iteration Methods in Electromagnetic Scattering* (De Gruyter, 2001).
155. J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley, 1994).
156. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, 1960).
157. A. Welters, personal communication.
158. A. F. Koenderink, “On the use of Purcell factors for plasmon antennas,” *Opt. Lett.* **35**, 4208–4210 (2010).
159. C. Sauvan, J. P. Hugonin, I. S. Maksymov, and P. Lalanne, “Theory of the spontaneous optical emission of nanosize photonic and plasmon resonators,” *Phys. Rev. Lett.* **110**, 237401 (2013).
160. A. Raman and S. Fan, “Photonic band structure of dispersive metamaterials formulated as a Hermitian eigenvalue problem,” *Phys. Rev. Lett.* **104**, 087401 (2010).
161. A. W. Snyder and J. Love, *Optical Waveguide Theory* (Kluwer Academic, 1983).
162. R. E. Collin, *Field Theory of Guided Waves* (IEEE, 1991), 2nd ed.
163. P. Leung, S. Liu, and K. Young, “Completeness and orthogonality of quasinormal modes in leaky optical cavities,” *Phys. Rev. A* **49**, 3057–3067 (1994).
164. P. T. Kristensen, C. Van Vlack, and S. Hughes, “Generalized effective mode volume for leaky optical cavities,” *Opt. Lett.* **37**, 1649–1651 (2012).
165. R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University, 1986).
166. C. E. Baum, E. J. Rothwell, K.-M. Chen, and D. P. Nyquist, “The singularity expansion method and its application to target identification,” *Proc. IEEE* **79**, 1481–1492 (1991).
167. C. E. Baum, “The singularity expansion method,” in “Transient Electromagnetic Fields,” (Springer, 1976), pp. 129–179.
168. A. Ramm, “Theoretical and practical aspects of singularity and eigenmode expansion methods,” *IEEE Trans. Antennas Propag.* **28**, 897–901 (1980).
169. C. L. Dolph and S. K. Cho, “On the relationship between the singularity expansion method and the mathematical theory of scattering,” *IEEE Trans. Antennas Propag.* **28**, 888–897 (1980).
170. A. M. van den Brink and K. Young, “Jordan blocks and generalized bi-orthogonal bases: realizations in open wave systems,” *J. Phys. A: Math. Gen.* **34**, 2607–2624 (2001).
171. M. Liertzer, L. Ge, A. Cerjan, A. D. Stone, H. E. Türeci, and S. Rotter, “Pump-induced exceptional points in lasers,” *Phys. Rev. Lett.* **108**, 173901 (2012).
172. M. Costabel, E. Darrigrand, and H. Sakly, “The essential spectrum of the volume integral operator in electromagnetic scattering by a homogeneous body,” *Comptes Rendus Math.* **350**, 193–197 (2012).
173. C. Loo, A. Lowery, N. Halas, J. West, and R. Drezek, “Immunotargeted nanoshells for integrated cancer imaging and therapy,” *Nano Lett.* **5**, 709–711 (2005).
174. S. Lal, S. Link, and N. J. Halas, “Nano-optics from sensing to waveguiding,” *Nat. Photonics* **1**, 641–648 (2007).
175. P. K. Jain, K. S. Lee, I. H. El-Sayed, and M. A. El-Sayed, “Calculated absorption and scattering properties of gold nanoparticles of different size, shape, and composition: applications in biological imaging and biomedicine,” *J. Phys. Chem. B* **110**, 7238–7248 (2006).
176. W. L. Barnes, A. Dereux, and T. W. Ebbesen, “Surface plasmon subwavelength optics,” *Nature* **424**, 824–830 (2003).
177. J. A. Dionne, L. A. Sweatlock, H. A. Atwater, and A. Polman, “Planar metal plasmon waveguides: frequency-dependent dispersion, propagation, localization, and loss beyond the free electron model,” *Phys. Rev. B* **72**, 1–11 (2005).
178. W. L. Barnes, “Surface plasmon-polariton length scales: a route to sub-wavelength optics,” *J. Opt. A Pure Appl. Opt.* **8**, S87–S93 (2006).
179. J. Homola, “Electromagnetic theory of surface plasmons,” in “Surface Plasmon Resonance Based Sensors,” J. Homola, ed. (Springer-Verlag, 2006), chap. 1, pp. 3–44.

## 1. Introduction

At optical frequencies, metals present a tradeoff: their conduction electrons enable highly subwavelength resonances, but at the expense of potentially significant electron-scattering losses [1–10]. In this article we formalize the tradeoff between resonant enhancement and loss, deriving limits to the absorption within, the scattering by, and the local density of states (LDOS) [11–16] near a lossy, absorptive body of arbitrary shape. Given a material of susceptibility  $\chi(\omega)$ , the limits depend only on a material enhancement factor  $|\chi(\omega)|^2/\text{Im}\chi(\omega)$  and on the incident-beam energy density (leading to a potential  $1/d^3$  LDOS enhancement for a metal–

emitter separation  $d$ ). The power scattered or dissipated by a material body must be smaller than the total power it extracts from an incident beam; we show that this statement of energy conservation yields limits to the magnitudes of the internal fields and polarization currents that control the scattering properties of a body. Unlike previous bounds [6, 9, 17–27], our limits do not depend on shape, size, or topology, nor do they diverge for zero bandwidth. The crucial ingredient is that our bounds depend on  $\chi$  and are finite only for realistic lossy materials—for idealized lossless materials such as perfect conductors, arbitrarily large optical responses are possible. We provide examples of structures that approach the theoretical limits, and also specific frequency ranges at which common structures fall far short. Our bounds apply to any absorptive system, and thus provide benchmarks for the response of metals, synthetic plasmonic materials (doped semiconductors) [3, 5, 28–30], and surface-phonon-polariton materials across visible and infrared wavelengths, resolving a fundamental question [1–10] about the extent to which resonant enhancement can overcome intrinsic dissipation.

There has been intense interest in exploiting “plasmonic” [2] effects, which arise for materials with permittivities that have negative real parts, in metals at optical frequencies. Geometries incorporating such materials are capable of supporting highly subwavelength surface resonances [1, 2]. Yet such a material has inherent loss arising from the typically significant imaginary part of  $\chi$ . Even for applications in which absorption is the goal, material loss dampens resonant excitations and reduces the overall response. This tradeoff between resonant enhancement and absorption has been investigated for specific geometries amenable to semianalytical methods, leading to a variety of geometry-dependent material dependences. For example, in the quasistatic limit, coated spheres absorb energy at a rate proportional to  $|\chi|/\text{Im}\chi$  (Ref. [31]), whereas spheroids absorb energy at a rate [32] proportional to  $|\chi|^2/\text{Im}\chi$ . Surface modes at planar metal–insulator interfaces exhibit propagation lengths proportional to  $(\text{Re}\epsilon)^2/\text{Im}\epsilon$  at very low frequencies [33] (in the Sommerfeld–Zenneck regime [2]), but near their surface-plasmon frequencies their propagation lengths are approximately proportional to  $\sqrt{\text{Im}\epsilon}$  (cf. App. E). In electron energy loss spectroscopy (EELS), the electron scattering cross-section is proportional to a “loss function”  $\text{Im}(-1/\epsilon) = (\text{Im}\epsilon)/|\epsilon|^2$  that enables experimental measurement of bulk plasmon frequencies [34, 35].

We show that  $|\chi|^2/\text{Im}\chi$  is a universal criterion for evaluating the optimal response of a metal, with a suitable generalization in Sec. 3 for more general media that may be anisotropic, magnetic, chiral, or inhomogeneous. For most materials,  $|\chi|^2/\text{Im}\chi$  increases as a function of wavelength (as demonstrated in Sec. 3.2), suggesting that if an optimal structure is known, the plasmonic response of a metal can potentially be much greater *away* from its bulk- and surface-plasmon frequencies. For effective-medium metamaterials, our bounds apply to both the underlying material parameters as well as to the effective medium parameters, with the *smaller* bound of the two controlling the maximum response. Thus effective-medium approaches cannot circumvent the bounds arising from their constitutive materials; however, they may find practical application if they can achieve resonances at frequencies that are otherwise difficult to achieve with the individual materials.

The limits derived here arise from basic energy considerations. An incident field  $\mathbf{E}_{\text{inc}}$  interacting with a scatterer generates polarization currents  $\mathbf{P}$  that depend on both the incident field and on the shape and susceptibility of the body. A lossy scatterer dissipates energy at a rate proportional to the squared magnitude of the currents,  $|\mathbf{P}|^2$ . At the same time, the total power extracted from the incident beam, i.e. the “extinction” (absorption plus scattering), is proportional to the imaginary part of the overlap integral of the polarization currents with the incident field,  $\sim \int_V \mathbf{E}_{\text{inc}}^* \cdot \mathbf{P}$ , which is known as the electromagnetic optical theorem [32, 36–38] and can be understood physically as the work done by the incident field to drive the induced currents. The overlap integral is only linear in  $\mathbf{P}$  whereas the absorption depends quadratically



on  $\mathbf{P}$ . If the magnitude of  $\mathbf{P}$  could increase without bound, then, the absorption would become greater than extinction, resulting in a physically impossible negative scattered power. Instead, there is a limit to the magnitude of the polarization field, and therefore to the scattering properties of any body comprising the lossy material. We make this argument precise in Sec. 3, where we employ variational calculus to derive general limits for a wide class of materials, and we also present limits specific to metals, which are typically homogeneous, isotropic, and nonmagnetic at optical frequencies. We consider here only bulk susceptibilities, excluding non-local or quantum effects [39–41]. The key results are the limits to absorption and scattering in Eqs. (29a,29b,31a–32b) and the limits to LDOS enhancement in Eqs. (34a,34b). Before deriving the limits, we present the volume-integral expressions for absorption, scattering, and radiative and nonradiative LDOS in Sec. 2.

In Sec. 4 we compare the response of a number of structures towards achieving the various limits. For far-field absorption and scattering, we find that ellipsoidal nanoparticles are ideal and can reach the limits across a wide range of frequencies. For near-field enhancement of power expended by a dipole into radiation or absorption, we find that it is much more difficult to reach the limits. The nonradiative LDOS near a planar metal surface reaches the limit at the “surface-plasmon frequency” of the metal. At lower frequencies, common structures (thin films, metamaterials) fall far short. At all frequencies, common designs for radiative LDOS enhancement exhibit suboptimal response. These results suggest the possibility for significant design improvement if the limits are achievable (i.e. “tight”).

Previous limits to the electromagnetic response of metals have emphasized a variety of limiting factors. At RF and millimeter-wave frequencies, where the response can be bounded relative to that of a perfect electric conductor (PEC) [42], the Wheeler–Chu–McLean limit [23–25, 42] bounds the radiative  $Q$  factor of an electrically small antenna. At optical frequencies, absorption loss increases and often dominates relative to radiative loss. There are known lower bounds on the absorptive  $Q$  for low-loss, quasistatic structures [26], and more generally for metals of any size with susceptibilities comprising Lorentz–Drude oscillator terms [9].

Limits to frequency-integrated extinction are also known. Purcell derived the first such limit, using the Kramers-Kronig relations [38] to bound the integrated response of spheroidal particles to their electrostatic ( $\omega = 0$ ) induced dipole moments [17]. Recently the limits have been extended to arbitrary shapes [18–20], but one obtains a different limit for each shape. Moreover, it is important in many applications to disentangle the single-frequency response and the bandwidth, and to do so separately for absorption and scattering.

Single-frequency absorption and scattering limits have primarily been derived via spherical-harmonic decompositions, originally for spherically symmetric scatterers [43, 44] and later for generic ones [21, 22]. This approach has been generalized recently, yielding limits in terms of the inverse of a scattered-field-operator [27], although it the inverse of such an operator is seemingly difficult to bound without resorting to spherical harmonics. In Sec. 3 we show that the scattered-field-operator approach and our material-dissipation approach share a common origin in volume-integral equations. The key distinction is that the scattered-field operator and its corresponding limits are independent of material but dependent on structure, whereas our limits incorporate material properties and are *independent* of structure. Both classes of limits apply to any linear body. In Sec. 5 we provide a more detailed comparison, finding that the spherical-harmonic limits may provide better design criteria at lower (e.g. rf) frequencies, whereas our limits should guide design at higher frequencies, especially in the field of plasmonics.

This work was partly inspired by our recent bounds [45] on extinction by quasistatic nanostructures. In that work we derived bounds via sum rules of quasistatic surface-integral operators [46, 47]; equivalently, we could have [48] derived the bounds via analogous constraints in composite theory [49, 50]. The key distinction between this work and our previous work [45] is

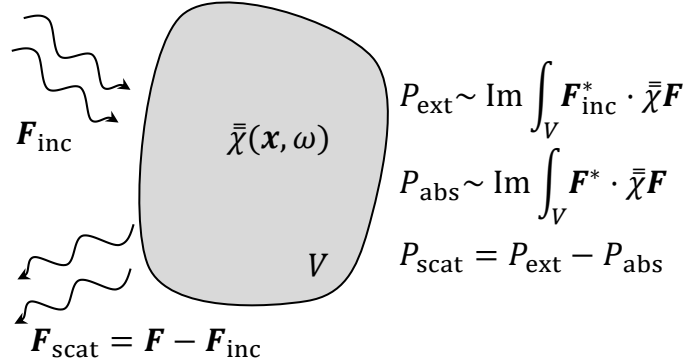


Fig. 1. Scattering problem under consideration. An incident field  $\mathbf{F}_{\text{inc}} = (\mathbf{E}_{\text{inc}} \quad Z_0 \mathbf{H}_{\text{inc}})^T$  impinges on a lossy scatterer with a susceptibility tensor  $\bar{\chi}(\mathbf{x}, \omega)$ . The shape and topology of the scatterer are arbitrary: It may be periodic, extend to infinity, or consist of multiple particles. The limits presented in Sec. 3 hinge on the fact that absorption is a quadratic functional of the electromagnetic fields, whereas extinction, by the optical theorem, is the imaginary part of a linear functional of the fields. In Sec. 3, we present general limits for tensor susceptibilities and also simplified limits for metals.

that here we find limits in the full Maxwell regime, such that our bounds apply to any structure at any size scale, and they apply to functions of the scattered fields (e.g. scattered power and radiative LDOS), which have zero amplitude in quasistatic electromagnetism. An additional benefit of our simplified energy-conservation approach is that we can bound the responses of anisotropic, magnetic, and/or inhomogeneous media, whereas the surface-integral sum-rule approach only works for isotropic and nonmagnetic materials. In this work we also consider the local density of states, which we did not consider previously and which represents an important design application.

## 2. Absorption, scattering, and LDOS expressions

We consider lossy media interacting with electromagnetic fields incident from fixed external sources (e.g. plane waves or dipole sources). Figure 1 illustrates the conceptual setup: a generic scatterer with susceptibility tensor  $\bar{\chi}$  absorbs, scatters, and extinguishes (extinction defined as absorbed plus scattered power) incident radiation at rates proportional to volume integrals over the scatterer. In this section we present the known volume-integral expressions for absorption and scattering, and we also derive volume-integral expressions for the power expended by a dipole near such a scatterer, which is either radiated to the far field or absorbed in the near field. Relative to a dipole in free space, the enhancement in power expended is given by the relative increase in the local density of states (LDOS) [15].

The scatterer is taken to consist of a lossy, local material that is possibly inhomogeneous, electric, magnetic, anisotropic, or bianisotropic (chiral). We assume the scatterer is in vacuum, with permittivity  $\epsilon_0$ , permeability  $\mu_0$ , impedance  $Z_0 = \sqrt{\mu_0/\epsilon_0}$ , and speed of light  $c = 1/\sqrt{\epsilon_0\mu_0}$ . Extending the limits of the following section to non-vacuum and possibly inhomogeneous backgrounds is relatively straightforward and is discussed in Sec. 5. The response of the scatterer can be described by induced electric and magnetic polarization currents,  $\mathbf{P}(\mathbf{x})$  and  $\mathbf{M}(\mathbf{x})$ , which satisfy the constitutive field relations

$$\begin{aligned} \mathbf{D}(\mathbf{x}) &= \epsilon_0 \mathbf{E}(\mathbf{x}) + \mathbf{P}(\mathbf{x}) \\ \mathbf{B}(\mathbf{x}) &= \mu_0 [\mathbf{H}(\mathbf{x}) + \mathbf{M}(\mathbf{x})]. \end{aligned} \quad (1)$$

For the general class of materials considered here, the currents  $\mathbf{P}$  and  $\mathbf{M}$  each depend on both  $\mathbf{E}$  and  $\mathbf{H}$  through a unitless  $6 \times 6$  susceptibility tensor  $\bar{\chi}$  [15, 51]:

$$\begin{pmatrix} \mathbf{P} \\ \frac{1}{c}\mathbf{M} \end{pmatrix} = \epsilon_0 \bar{\chi} \begin{pmatrix} \mathbf{E} \\ Z_0 \mathbf{H} \end{pmatrix} = \epsilon_0 \bar{\chi} \mathbf{F} \quad (2)$$

where  $\mathbf{F}$  is a generalized vector field containing both electric and magnetic fields. For isotropic media with relative permittivity  $\epsilon_r$  and relative permeability  $\mu_r$ , the susceptibility tensor comprises only the diagonal elements  $\epsilon_r - 1$  and  $\mu_r - 1$ . Lossy media as considered here have susceptibilities that satisfy the positive-definiteness condition (for an  $e^{-i\omega t}$  time convention) [51, 52]

$$\omega \text{Im} \bar{\chi} > 0, \quad (3)$$

where  $\text{Im} \bar{\chi} = (\bar{\chi} - \bar{\chi}^\dagger)/2i$  (and  $\dagger$  represents the conjugate transpose).

When light impinges on the scatterer, the absorbed, scattered, and extinguished powers can be written as overlap integrals of the internal currents and fields [53]. The absorption (dissipation) within such a medium is the work done by the total fields on the induced currents, given by the expression [54]:

$$P_{\text{abs}} = \frac{\epsilon_0 \omega}{2} \text{Im} \int_V \mathbf{F}^*(\mathbf{x}) \cdot \bar{\chi}(\mathbf{x}) \mathbf{F}(\mathbf{x}) dV \quad (4)$$

where the asterisk denotes complex conjugation. Equation (4) reduces to the usual  $(\epsilon_0 \omega/2) (\text{Im} \chi) \int_V |\mathbf{E}|^2$  for homogeneous, isotropic, nonmagnetic media.

The total power extracted from the incident fields—the extinction—is the sum of the absorbed and scattered powers and can be computed by the optical theorem [38, 55]. Although commonly written as an integral over fictitious effective surface currents [38], the optical theorem can also be written as a volume integral over the polarization currents [36, 53, 56], representing the work done by the *incident* fields on the induced currents:

$$P_{\text{ext}} = \frac{\epsilon_0 \omega}{2} \text{Im} \int_V \mathbf{F}_{\text{inc}}^*(\mathbf{x}) \cdot \bar{\chi}(\mathbf{x}) \mathbf{F}(\mathbf{x}) dV \quad (5)$$

where, as for  $\mathbf{F}$ , we define  $\mathbf{F}_{\text{inc}}$  by

$$\mathbf{F}_{\text{inc}} = \begin{pmatrix} \mathbf{E}_{\text{inc}} \\ Z_0 \mathbf{H}_{\text{inc}} \end{pmatrix}. \quad (6)$$

The scattered power is the difference between extinction and absorption:

$$\begin{aligned} P_{\text{scat}} &= P_{\text{ext}} - P_{\text{abs}} \\ &= \frac{\epsilon_0 \omega}{2} \text{Im} \int_V [\mathbf{F}_{\text{inc}}^*(\mathbf{x}) - \mathbf{F}^*(\mathbf{x})] \cdot \bar{\chi}(\mathbf{x}) \mathbf{F}(\mathbf{x}) dV. \end{aligned} \quad (7)$$

In addition to absorbing and scattering light, structured media can also alter the spontaneous emission rates of nearby emitters. Increased spontaneous emission shows exciting potential for surface-enhanced Raman scattering (SERS) [57, 58], fluorescent imaging [59, 60], thermophotovoltaics [61, 62], and ultrafast light-emitting diodes (LEDs) [63]. The common metric for the enhanced emission rate is the (electric) local density of states (LDOS), which represents the density of modes weighted by the relative energy density of each mode's electric field at a given position [15]. Equivalently, and more generally, the LDOS enhancement represents the enhancement in the total power expended by an electric dipole radiator [15, 64, 65], into either

radiation or dissipation. Similar to extinction, the total LDOS is the imaginary part of a linear functional of the induced electric fields [12]:

$$\rho_{\text{tot}} = \frac{1}{\pi\omega} \text{Im} \sum_j \hat{\mathbf{s}}_j \cdot \mathbf{E}_{s_j}(\mathbf{x}_0), \quad (8)$$

where  $\mathbf{E}_{s_j}$  denotes the field from a dipole source at  $\mathbf{x}_0$  polarized in the  $\hat{\mathbf{s}}_j$  direction, with a dipole moment  $\mathbf{p}_0 = \epsilon_0 \hat{\mathbf{s}}_j$ , and the sum over  $j$  accounts for all possible orientations (the conventional LDOS corresponds to a randomly oriented dipole [12]).

To connect the LDOS to the material properties, we rewrite it as a volume integral over the fields within the scatterer. The total field at the source position,  $\mathbf{E}_{s_j}(\mathbf{x}_0)$ , consists of an incident field and a scattered field:

$$\mathbf{E}_{s_j}(\mathbf{x}_0) = \mathbf{E}_{\text{inc},s_j}(\mathbf{x}_0) + \mathbf{E}_{\text{scat},s_j}(\mathbf{x}_0). \quad (9)$$

The incident field is known—it is the field of a dipole in vacuum—and can be left as-is (note that the imaginary part of a dipole field does not diverge at the source location [66]). The scattered field arises from interactions with the scatterer and is the composite field from the induced electric and magnetic currents, radiating as if in free space:

$$\mathbf{E}_{\text{scat}}(\mathbf{x}_0) = \int_V \left[ \overline{\overline{G^{EP}}}(\mathbf{x}_0, \mathbf{x}) \mathbf{P}(\mathbf{x}) + \overline{\overline{G^{EM}}}(\mathbf{x}_0, \mathbf{x}) \mathbf{M}(\mathbf{x}) \right] \quad (10)$$

where  $\overline{\overline{G^{EP}}}$  and  $\overline{\overline{G^{EM}}}$  are free-space dyadic Green's functions [51, 67]. We could at this point insert Eqs. (9,10) into Eq. (8) and have a volume-integral equation for the total LDOS. However, note that the Green's functions in Eq. (10) represent the fields of free-space dipoles and the excitation in this case is also a dipole. We follow this intuition to replace the Green's functions by the incident fields.

Whereas the source dipole at  $\mathbf{x}_0$  generates incident fields at points  $\mathbf{x}$  within the scatterer, the Green's functions in Eq. (10) are the fields at  $\mathbf{x}_0$  from a dipole at  $\mathbf{x}$ . By reciprocity [68] one can switch the source and destination points of vacuum Green's functions

$$G_{ij}^{EP}(\mathbf{x}_0, \mathbf{x}) = G_{ji}^{EP}(\mathbf{x}, \mathbf{x}_0) \quad (11a)$$

$$G_{ij}^{EM}(\mathbf{x}_0, \mathbf{x}) = -\mu_0 G_{ji}^{HP}(\mathbf{x}, \mathbf{x}_0) \quad (11b)$$

where for clarity we indexed the Green's function tensors. Now one can see that the product  $\hat{\mathbf{s}}_j \cdot \overline{\overline{G^{EP}}} \mathbf{P}$  equals  $\epsilon_0^{-1} \mathbf{E}_{\text{inc}} \cdot \mathbf{P}$ . The magnetic Green's function yields the incident magnetic field, with a negative sign arising from reciprocity, as in Eq. (11b). Equation (9) can be written

$$\hat{\mathbf{s}}_j \cdot \mathbf{E}_{s_j}(\mathbf{x}_0) = \hat{\mathbf{s}}_j \cdot \mathbf{E}_{\text{inc},s_j}(\mathbf{x}_0) + \int_V \tilde{\mathbf{F}}_{\text{inc},s_j} \cdot \overline{\overline{\chi}} \mathbf{F}_{s_j} \quad (12)$$

where we have defined

$$\tilde{\mathbf{F}}_{\text{inc}} = \begin{pmatrix} \mathbf{E}_{\text{inc}} \\ -Z_0 \mathbf{H}_{\text{inc}} \end{pmatrix}. \quad (13)$$

Inserting the field equation, Eq. (12), into the LDOS equation, Eq. (8), yields the total LDOS as the sum of free-space and scattered-field contributions:

$$\rho_{\text{tot}} = \rho_0 + \frac{1}{\pi\omega} \text{Im} \sum_j \int_V \tilde{\mathbf{F}}_{\text{inc},s_j} \cdot \overline{\overline{\chi}} \mathbf{F}_{s_j}. \quad (14)$$

where  $\rho_0 = \omega^2/2\pi^2c^3$  is the free-space electric LDOS [69]. It is typically more useful to normalize  $\rho$  to  $\rho_0$ :

$$\frac{\rho_{\text{tot}}}{\rho_0} = 1 + \frac{2\pi}{k^3} \text{Im} \sum_j \int_V \tilde{\mathbf{F}}_{\text{inc},s_j} \cdot \bar{\chi} \mathbf{F}_{s_j} \quad (15)$$

where  $k = \omega/c$ . Equation (15) relates the total LDOS to a volume integral over the scatterer, which will enable us to find upper bounds to the response in the next section.

For many applications it is important to distinguish between the power radiated by the dipole into the far-field (where it may be imaged, for example) and the power absorbed in the near field (which may productively transfer heat, for example). Absorbed power is given by Eq. (4), and thus the nonradiative LDOS enhancement  $\rho_{\text{nr}}/\rho_0$  is given by Eq. (4) divided by the power radiated by a dipole (of amplitude  $\epsilon_0$ ) in free space,  $P_{\text{rad}} = \epsilon_0\omega^4/12\pi c^3$  (Ref. [38]):

$$\frac{\rho_{\text{nr}}}{\rho_0} = \frac{2\pi}{k^3} \text{Im} \sum_j \int_V \mathbf{F}_{s_j}^* \cdot \bar{\chi} \mathbf{F}_{s_j}. \quad (16)$$

Finally, just as the scattered power in Eq. (7) is the difference between extinction and absorption, the radiative part of the LDOS is the difference between the total and nonradiative parts:

$$\frac{\rho_{\text{rad}}}{\rho_0} = 1 + \frac{2\pi}{k^3} \text{Im} \sum_j \int_V [\tilde{\mathbf{F}}_{\text{inc},s_j} - \mathbf{F}_{s_j}^*] \cdot \bar{\chi} \mathbf{F}_{s_j} \quad (17)$$

### 3. Limits

Given the power and LDOS expressions of the previous section, upper bounds to each quantity can be derived by exploiting the energy conservation ideas discussed in the introduction (Sec. 1). The extinction is the imaginary part of a linear function of the polarization currents, whereas absorption is proportional to their squared magnitude (and scattered power is the difference between the two), and thus energy conservation yields finite optimal polarization currents and fields for each quantity.

Just as one can use gradients to find stationary points in finite-dimensional calculus, one can use variational derivatives [70] to find stationary points of a *functional* (i.e. a function of a function). It is sufficient here to consider functionals of the type  $P = \int f^*g$ , which arise in the power expressions, Eqs. (4,7,16,17). The variational derivative of  $P$  with respect to  $g$  is given by  $\frac{\delta}{\delta g} \int f^*g = f^*$ , analogous to the gradient in vector calculus:  $\nabla_{\mathbf{x}} (\mathbf{a}^\dagger \mathbf{x}) = \mathbf{a}^\dagger$ . The primary distinction is the dimensionality of the space and thus the appropriate choice of inner product.

The optimal fields for the various response functions  $P$  are therefore those for which  $P$  is stationary under small variations of the field degrees of freedom. The field  $\mathbf{F}$  is complex-valued, such that one could take the real and imaginary parts of  $\mathbf{F}$  as independent ( $P$  is a nonconstant real-valued functional and therefore not analytic [71] in  $\mathbf{F}$ ), but a more natural choice is to formally treat the field  $\mathbf{F}$  and its complex-conjugate  $\mathbf{F}^*$  as independent variables [71, 72]. Then a necessary condition for an extremum of a functional  $P[\mathbf{F}]$  is for the variational derivatives with respect to the field degrees of freedom to equal zero,  $\delta P/\delta \mathbf{F} = 0$  and  $\delta P/\delta \mathbf{F}^* = 0$  (which are the Euler-Lagrange equations [70] for functionals that do not depend on the gradients of their arguments). Because our response functions are real-valued, the derivatives with respect to  $\mathbf{F}$  and  $\mathbf{F}^*$  are redundant—they are complex-conjugates of each other [71]—and the condition for the extremum can be found with the single equation

$$\frac{\delta P}{\delta \mathbf{F}^*} = 0, \quad (18)$$

where we have chosen to vary  $\mathbf{F}^*$  instead of  $\mathbf{F}$  for its slightly simpler notation going forward. We apply this variational calculus approach to bound each response function of interest. First, we derive limits for the most general class of materials under consideration. Then we specialize to metals, an important class of lossy media that are typically homogeneous, isotropic, and nonmagnetic at optical frequencies.

### 3.1. General lossy media

We consider first Eq. (7), for the scattered power. Setting the variational derivative of  $P_{\text{scat}}$  to 0 yields

$$\frac{\delta P_{\text{scat}}}{\delta \mathbf{F}^*} = -\frac{\epsilon_0 \omega}{2} \left[ \frac{\overline{\chi}^\dagger \mathbf{F}_{\text{inc}}}{2i} + (\text{Im} \overline{\chi}) \mathbf{F} \right] = 0. \quad (19)$$

The optimal field that satisfies Eq. (19) is

$$\mathbf{F}_{\text{scat,opt}}(\mathbf{x}) = \frac{i}{2} [\text{Im} \overline{\chi}(\mathbf{x})]^{-1} \overline{\chi}^\dagger(\mathbf{x}) \mathbf{F}_{\text{inc}}(\mathbf{x}) \quad (20)$$

for all points  $\mathbf{x}$  within the scatterer volume  $V$ . The optimal field is guaranteed to exist because  $\omega \text{Im} \overline{\chi}$  is positive-definite, per Eq. (3), and therefore invertible. We have only shown that this is an extremum, not a maximum, but because  $\omega \text{Im} \overline{\chi}$  is positive-definite, the scattered power in Eq. (7) is a concave functional, for which any extremum must be a global maximum [73].

One can see that the optimal fields within the scatterer are related to the incident fields (directly proportional for homogeneous media), which conforms intuitively with the scattered-power expression in Eq. (7). The internal field should strongly overlap with the incident field, to increase the power extracted from the incident beam, while the susceptibility dependence balances between maximizing extinction and minimizing absorption.

A similar procedure yields the optimal internal fields for maximum absorption within a scatterer. Although the absorbed power as given in Eq. (4) is unbounded with respect to  $\mathbf{F}$ , adding the constraint that absorption must be smaller than extinction (i.e. the scattered power must be nonnegative) imposes an upper bound. Because Eq. (4) is unbounded, the Karash–Kuhn–Tucker (KKT) conditions [74] require that the constraint  $P_{\text{scat}} \geq 0$  must be *active*, i.e.  $P_{\text{scat}} = 0$ . Following standard constrained-optimization theory [74], we define the Lagrange multiplier  $\ell$  and the Lagrangian functional  $\mathcal{L} = P_{\text{abs}} + \ell P_{\text{scat}}$ . The extrema of  $\mathcal{L}$  satisfy

$$\begin{aligned} \frac{\delta \mathcal{L}}{\delta \mathbf{F}^*} &= \frac{\delta P_{\text{abs}}}{\delta \mathbf{F}^*} + \ell \frac{\delta P_{\text{scat}}}{\delta \mathbf{F}^*} \\ &= \frac{\epsilon_0 \omega}{2} \left[ (1 - \ell) (\text{Im} \overline{\chi}) \mathbf{F} - \frac{\ell}{2i} \overline{\chi}^\dagger \mathbf{F}_{\text{inc}} \right] = 0. \end{aligned} \quad (21)$$

To simultaneously ensure that the scattered power also equals 0, one can verify that the Lagrange multiplier is given by  $\ell = 2$ . Then the optimal internal fields are

$$\mathbf{F}_{\text{abs,opt}}(\mathbf{x}) = i [\text{Im} \overline{\chi}(\mathbf{x})]^{-1} \overline{\chi}^\dagger(\mathbf{x}) \mathbf{F}_{\text{inc}}(\mathbf{x}) \quad (22)$$

which are precisely double the optimal scattering fields of Eq. (20). Maximizing  $P_{\text{abs}}$  subject to  $P_{\text{scat}} \geq 0$  is a problem of maximizing a convex functional subject to a convex quadratic constraint, such that the solution in Eq. (22) must be a global maximum [75].

The limits to the scattered and absorbed powers are given by substituting the optimal fields

in Eq. (20) and Eq. (22) into Eq. (7) and Eq. (4), respectively:

$$P_{\text{scat}} \leq \frac{\epsilon_0 \omega}{8} \int_V \mathbf{F}_{\text{inc}}^* \cdot \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \mathbf{F}_{\text{inc}} d^3 \mathbf{x} \quad (23a)$$

$$P_{\text{abs}}, P_{\text{ext}} \leq \frac{\epsilon_0 \omega}{2} \int_V \mathbf{F}_{\text{inc}}^* \cdot \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \mathbf{F}_{\text{inc}} d^3 \mathbf{x} \quad (23b)$$

where extinction has the same limit as absorption, which can be derived by maximizing  $P_{\text{ext}}$  subject to  $P_{\text{scat}} \geq 0$ . The limits depend only on the intensity of the incident field and the material susceptibility  $\bar{\chi}$  over the volume of the scatterer. The product  $\bar{\chi} (\text{Im} \bar{\chi})^{-1} \bar{\chi}^\dagger$ , discussed further below, sets the bound on how large the induced currents can be in a dissipative medium. Whereas optimal per-volume scattering occurs under a condition of equal absorption and scattering, optimal per-volume absorption occurs in the absence of any scattered power and can be larger by a factor of four. This ordering is reversed in the spherical-multipole limits [21, 22], where the scattering cross-section (*not* normalized by volume) can be four times larger than the absorption cross-section.

Note that Eqs. (23a,23b) look superficially similar to the absorbed- and scattered-power limits in Ref. [27]. As discussed in Sec. 1, they share a common origin as energy-conservation principles applied to integral equations. The key distinction is which quantity serves as a non-negative quadratic (in  $\mathbf{F}$ ) constraint. We treat absorption as the quadratic quantity, given by Eq. (4), with the scattered power as the difference between extinction and absorption. The scattered-field-operator approach rewrites the scattered power via a volume integral equation (VIE) [51]. This yields a non-negative, quadratic scattered power that is of the same form as Eq. (4) except with the replacement  $\text{Im} \bar{\chi} \rightarrow \text{Im} \mathcal{G}$ , where  $\mathcal{G}$  is a scattered-field integral operator with the homogeneous Green's function as its kernel (the electric component of  $\mathcal{G}$  is used in Eq. (10)). Energy conservation leads to the limits of Ref. [27], which take a similar form to Eqs. (23a,23b), except  $\text{Im} \bar{\chi} \rightarrow \text{Im} \mathcal{G}$  and the factors of four are reversed. The limits in Ref. [27] are a generalization of the spherical-multipole limits of antenna theory [21, 22], which treat the special case of the scattered field decomposed into spherical harmonics.

Our limits have very different characteristics from those of [21, 22, 27]. Our approach, via absorption as the quadratic constraint, yields limits that incorporate the material properties and are independent of structure. This naturally results in per-volume limits, a normalization of inherent interest to designers. The scattered-field-operator approach yields limits that are independent of material but depend on the structure, in a way that can be difficult to be quantified because the inverse of the scattered-field operator is not known except for the simplest cases (e.g. dipoles). A spherical-harmonic decomposition of the operator yields analytical limits, but only to the cross-sections, without normalization. The cross-section is inherently unbounded (increasing linearly with the geometric cross-section at large sizes) and thus difficult to use from a design perspective. The different normalizations are responsible for the different orderings of the absorbed- and scattered-power limits. In Sec. 5 we extend this comparison to show that our material-dissipation approach provides better design criteria at optical frequencies.

The same derivations lead to optimal fields and upper bounds for the radiative and nonradiative LDOS. The optimal fields are nearly identical in form:

$$\mathbf{F}_{s_j, \text{rad}, \text{opt}}(\mathbf{x}) = \frac{i}{2} (\text{Im} \bar{\chi})^{-1} \bar{\chi}^\dagger \tilde{\mathbf{F}}_{\text{inc}, s_j}^*(\mathbf{x}) \quad (24a)$$

$$\mathbf{F}_{s_j, \text{nr}, \text{opt}}(\mathbf{x}) = i (\text{Im} \bar{\chi})^{-1} \bar{\chi}^\dagger \tilde{\mathbf{F}}_{\text{inc}, s_j}^*(\mathbf{x}) \quad (24b)$$

where the complex conjugation arises due to the lack of conjugation in the LDOS expressions (which itself arises because open scattering problems in electromagnetism are complex-

symmetric rather than Hermitian [15]). Substituting the optimal fields into the LDOS expressions Eqs. (16,17) gives the LDOS limits

$$\frac{\rho_{\text{rad}}}{\rho_0} \leq 1 + \frac{\pi}{2k^3} \sum_j \int_V \tilde{\mathbf{F}}_{\text{inc},s_j} \cdot \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \tilde{\mathbf{F}}_{\text{inc},s_j}^* d^3 \mathbf{x} \quad (25a)$$

$$\frac{\rho_{\text{nr}}}{\rho_0}, \frac{\rho_{\text{tot}}}{\rho_0} \leq \frac{2\pi}{k^3} \sum_j \int_V \tilde{\mathbf{F}}_{\text{inc},s_j} \cdot \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \tilde{\mathbf{F}}_{\text{inc},s_j}^* d^3 \mathbf{x}. \quad (25b)$$

As for extinction, the limit to the total LDOS is identical to the limit to the nonradiative LDOS, which can be proven by maximizing  $\rho_{\text{tot}}$  subject to  $\rho_{\text{rad}} \geq 0$ .

The absorption, scattering, and LDOS limits in Eqs. (23a,23b,25a,25b) depend on the overlap integral of the material susceptibility and the incident field. We can simplify the limits further by separating the dependencies, which is simple for homogeneous, isotropic media but can also be done for more general media through induced matrix norms [76]. The integrand in Eq. (23a) (and each of the other limits) is of the form  $\mathbf{z}^\dagger \mathbf{A} \mathbf{z}$ , a quantity related to the *norm* (i.e. “magnitude”) of a matrix  $\mathbf{A}$ . The induced 2-norm of a matrix  $\mathbf{A}$ , denoted  $\|\mathbf{A}\|_2$ , is given by the maximum value of the quantity  $\mathbf{z}^\dagger \mathbf{A} \mathbf{z} / \mathbf{z}^\dagger \mathbf{z}$  for all  $\mathbf{z} \neq \mathbf{0}$ . The integral in Eq. (23a) can then be bounded for general media and arbitrary incident fields:

$$\begin{aligned} \int_V \mathbf{F}_{\text{inc}}^* \cdot \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \mathbf{F}_{\text{inc}} &\leq \int_V \left\| \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \right\|_2 \mathbf{F}_{\text{inc}}^* \cdot \mathbf{F}_{\text{inc}} \\ &\leq \left( \left\| \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \right\|_2 \right)_{\text{max}} \int_V |\mathbf{F}_{\text{inc}}|^2 \end{aligned} \quad (26)$$

where the dependence on the material susceptibility is now separated from the properties of the incident field  $\mathbf{F}_{\text{inc}}$ . The field intensity  $|\mathbf{F}_{\text{inc}}|^2$  is proportional to the energy density of the incident field:

$$\begin{aligned} \frac{1}{2} \epsilon_0 |\mathbf{F}_{\text{inc}}|^2 &= \frac{1}{2} \epsilon_0 |\mathbf{E}_{\text{inc}}|^2 + \frac{1}{2} \mu_0 |\mathbf{H}_{\text{inc}}|^2 \\ &= U_{E,\text{inc}} + U_{H,\text{inc}} \end{aligned} \quad (27)$$

where  $U_{E,\text{inc}}$  and  $U_{H,\text{inc}}$  are the (spatially varying) incident electric and magnetic energy densities [38]. Generally the incident fields relevant to  $P_{\text{scat}}$  and  $P_{\text{abs}}$  are beams with nearly constant intensity and infinite total energy, for which one should bound the scattered or absorbed power per unit volume of material. Given the operator definition and energy-density relation just discussed, the scattering and absorption limits in Eqs. (23a,23b) simplify:

$$\frac{P_{\text{scat}}}{V} \leq \frac{\omega (U_{E,\text{inc}} + U_{H,\text{inc}})_{\text{avg}}}{4} \left( \left\| \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \right\|_2 \right)_{\text{max}} \quad (28a)$$

$$\frac{P_{\text{abs}}}{V}, \frac{P_{\text{ext}}}{V} \leq \omega (U_{E,\text{inc}} + U_{H,\text{inc}})_{\text{avg}} \left( \left\| \bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi} \right\|_2 \right)_{\text{max}}. \quad (28b)$$

Plane waves are incident fields of general interest. They have equal electric and magnetic energy densities and constant intensities  $I_{\text{inc}} = cU_{E,\text{inc}}$ , where  $c$  is the speed of light in vacuum. The cross-section of a scatterer is defined  $\sigma = P/I_{\text{inc}}$ , representing the effective area the scatterer presents to the plane wave. Because plane waves are constant in space, the absorption and scattering bounds are tighter. In the first line of Eq. (26),  $|\mathbf{F}_{\text{inc}}|^2$  can be taken out of the integral, which then simplifies to the *average* value of the norm of  $\bar{\chi}^\dagger (\text{Im} \bar{\chi})^{-1} \bar{\chi}$ . With this modification



to Eqs. (28a,28b), the bounds on absorption and scattering cross-sections per unit volume are:

$$\frac{\sigma_{\text{scat}}}{V} \leq \frac{k}{2} \left( \left\| \bar{\chi}^\dagger (\text{Im } \bar{\chi})^{-1} \bar{\chi} \right\|_2 \right)_{\text{avg}} \quad (29a)$$

$$\frac{\sigma_{\text{abs}}}{V}, \frac{\sigma_{\text{ext}}}{V} \leq 2k \left( \left\| \bar{\chi}^\dagger (\text{Im } \bar{\chi})^{-1} \bar{\chi} \right\|_2 \right)_{\text{avg}}. \quad (29b)$$

which apply for general  $6 \times 6$  electric and magnetic susceptibility tensors. For susceptibilities that are only electric or only magnetic, and therefore  $3 \times 3$  tensors, the bound is smaller by a factor of two, since the incident magnetic field cannot drive magnetic currents (or vice versa). The LDOS analogue of Eqs. (29a,29b) is not straightforward, because the incident fields are inhomogeneous. Consequently, we leave Eqs. (25a,25b) as the general LDOS limits for inhomogeneous media, and derive a simpler version for metals in the next subsection.

Nanoparticle scattering and absorption are often written in terms of electric/magnetic polarizabilities and higher-order moments [22, 32, 43, 77], whereas Eqs. (29a,29b) are bounds in terms of only the material susceptibility. One implication is that Eqs. (29a,29b) imply restrictions on the number of moments that can be excited, or the strengths of the individual excitations, in a lossy scatterer. A lossy scatterer of finite size cannot have arbitrarily many spherical-multipole moments excited, nor can a single scatterer of very small size achieve full coupling to the lowest-order electric and magnetic dipole moments. Scatterers for which  $\text{Im } \chi / |\chi|^2 \gg V / \lambda^3$  cannot achieve  $\sim \lambda^2$  cross-sections per ‘‘channel,’’ even on resonance.

### 3.2. Metals

Metals represent an important and prevalent example of lossy media. (We define a material to behave as a ‘‘metal’’ at a given frequency  $\omega$  if  $\text{Re } \chi(\omega) < -1$ , thus including materials such as SiC [78, 79] and SiO<sub>2</sub> [79] that support surface-phonon polaritons at infrared wavelengths.) At optical frequencies, common metals have homogeneous, isotropic, and nonmagnetic susceptibilities, enabling us to write the matrix norm of the previous subsection as a simple scalar quantity,

$$\left\| \bar{\chi}^\dagger (\text{Im } \bar{\chi})^{-1} \bar{\chi} \right\|_2 = \frac{|\chi(\omega)|^2}{\text{Im } \chi(\omega)}, \quad (30)$$

where  $\chi(\omega)$  is the electric susceptibility. Another alteration in the metal case is that the incident magnetic energy density,  $U_{H,\text{inc}}$ , drops out of the limits because the magnetic polarization currents are zero in Eq. (2) (and therefore one can simplify  $\mathbf{F}_{\text{inc}}, \tilde{\mathbf{F}}_{\text{inc}} \rightarrow \mathbf{E}_{\text{inc}}$ ). This is not a quasistatic restriction to small objects that only interact with the incident electric field; larger objects that potentially interact strongly with the magnetic field remain valid. But their optimal response can be written in terms of only the incident electric field, since absorption and extinction by nonmagnetic objects can also be written only in terms of electric fields, per Eqs. (4,5).

The limits to per-volume absorption and scattering, simplifying Eqs. (28a,28b), are

$$\frac{P_{\text{scat}}}{V} \leq \frac{\omega U_{E,\text{inc}}}{4} \frac{|\chi(\omega)|^2}{\text{Im } \chi(\omega)} \quad (31a)$$

$$\frac{P_{\text{abs}}}{V} \leq \omega U_{E,\text{inc}} \frac{|\chi(\omega)|^2}{\text{Im } \chi(\omega)}. \quad (31b)$$

Similarly, the cross-section limits (reduced by a factor of two relative to Eqs. (29a,29b) because

$U_{H,\text{inc}}$  is not in the limit) are

$$\frac{\sigma_{\text{scat}}}{V} \leq \frac{k}{4} \frac{|\chi(\omega)|^2}{\text{Im} \chi(\omega)} \quad (32a)$$

$$\frac{\sigma_{\text{abs}}}{V}, \frac{\sigma_{\text{ext}}}{V} \leq k \frac{|\chi(\omega)|^2}{\text{Im} \chi(\omega)}, \quad (32b)$$

where as before  $k = \omega/c$ . Whereas the optimal per-volume scattering occurs at a condition of equal absorption and scattering, in App. C we also derive limits under a constraint of suppressed absorption, as may be desirable e.g. in a solar cell enhanced by plasmonic scattering [4].

The limits to the power expended by a nearby dipole emitter can be similarly simplified for metals. The incident field is the field of an electric dipole in free space, proportional to the product of the homogeneous Green's function and the dipole polarization vector,  $\mathbf{F}_{\text{inc},s_j} = \mathbf{E}_{\text{inc},s_j} = \overline{\overline{G^{EP}}} \hat{\mathbf{s}}_j$  (because the metal is nonmagnetic, only the incident electric field is relevant). The integral over the incident field in Eqs. (25a,25b), summed over dipole orientations, is given by  $\sum_j \int_V |\mathbf{E}_{\text{inc},s_j}|^2 = \int_V \|\overline{\overline{G^{EP}}}\|_F^2$ , where  $\|\cdot\|_F$  denotes the Frobenius norm [76]. For the homogeneous photon Green's function the squared Frobenius norm is shown in App. G to be

$$\|\overline{\overline{G^{EP}}}\|_F^2 = \frac{k^6}{8\pi^2} \left[ \frac{3}{(kr)^6} + \frac{1}{(kr)^4} + \frac{1}{(kr)^2} \right], \quad (33)$$

where the  $1/r^6$  and  $1/r^4$  terms arise from near-field nonradiative evanescent waves, and the  $1/r^2$  term corresponds to far-field radiative waves.

Inserting Eq. (33) into the LDOS limits, Eqs. (25a,25b), yields a complicated integral that depends on the exact shape of the body. The integrand is positive, though, so one can instead calculate a limit by integrating over a larger space that encloses the body (we will show that most of the potential for enhancement occurs very close to the emitter, such that the exact shape of the enclosure is usually irrelevant). We consider in detail the case in which the scatterer is contained within a half-space, but we also note immediately after Eqs. (34a,34b) the necessary coefficient replacement if the enclosure is a spherical shell. All structures separated from an emitter must fit into a spherical shell, and thus we have not imposed any structural restrictions (in particular, there is no need for a separating plane between the emitter and the scatterer).

We consider a finite-size approximation of the half-space: a circular cylinder enclosing the metal body, a distance  $d$  from the emitter and with equal height and radius,  $L$  (ultimately we are interested in the limit  $L \rightarrow \infty$ ). The volume integrals are straightforward in cylindrical coordinates, yielding  $\int_V 1/r^6 = \pi/6d^3$ ,  $\int_V 1/r^4 = \pi/d$ , and  $\int_V 1/r^2 = \pi \ln(2)L$ , for  $L \gg d$  (discarding the contributions  $\sim d/L$  for the evanescent-wave terms). Then the limits to radiative and non-radiative LDOS rates are

$$\frac{\rho_{\text{rad}}}{\rho_0} \leq \frac{|\chi(\omega)|^2}{\text{Im} \chi(\omega)} \left[ \frac{1}{32(kd)^3} + \frac{1}{16kd} + \text{O}(kL) \right] + 1 \quad (34a)$$

$$\frac{\rho_{\text{nr}}}{\rho_0} \leq \frac{|\chi(\omega)|^2}{\text{Im} \chi(\omega)} \left[ \frac{1}{8(kd)^3} + \frac{1}{4kd} + \text{O}(kL) \right] \quad (34b)$$

where  $\text{O}(\cdot)$  signifies ‘‘Big-O’’ notation [80]. Note that the  $\text{O}(kL)$  terms, which arise from the far-field excitation, diverge as the size  $L$  of the bounding region goes to  $\infty$ , whereas one would

expect the near-field excitation to be most important. The  $O(kL)$  divergence as  $L \rightarrow \infty$  is unphysical: it represents a polarization current that is proportional to the  $1/r$  incident field, according to Eqs. (24a,24b), over the entire half-space, maintaining a constant energy flux *within* a lossy medium. Hence, this  $O(kL)$  term, while a correct upper bound, is overly optimistic, and the attainable radiative contribution must be non-diverging in  $L$ . One could attempt to separately bound the evanescent and radiative excitations. However,  $L$  also represents the largest interaction distances over which polarization currents contribute to the LDOS (in Eq. (10), for example); in App. B we show that for reasonable interaction lengths  $L$  and near-field separations  $d$ , the contribution of the  $O(kL)$  terms is negligible compared to the  $1/d^3$  terms (because the divergence is slow). Thus in the near field, where the possibility for LDOS enhancement is most significant, the limits are dominated by the  $1/d^3$  terms:

$$\frac{\rho_{\text{rad}}}{\rho_0} \leq \frac{1}{32(kd)^3} \frac{|\chi(\omega)|^2}{\text{Im}\chi(\omega)} \quad (35a)$$

$$\frac{\rho_{\text{nr}}}{\rho_0}, \frac{\rho_{\text{tot}}}{\rho_0} \leq \frac{1}{8(kd)^3} \frac{|\chi(\omega)|^2}{\text{Im}\chi(\omega)}. \quad (35b)$$

A spherical-shell enclosure of solid angle  $\Omega$  yields the same result but with the replacement  $1/8 \rightarrow \Omega/4\pi$  in each limit. Again we see the possibility for enhancement proportional to  $|\chi|^2/\text{Im}\chi$ . There is the additional possibility of near-field enhancement proportional to  $1/(kd)^3$ , which arises from the increased amplitude of the incident field at the metal scatterer.

Figure 2 depicts  $|\chi|^2/\text{Im}\chi$  as a function of wavelength for many natural and synthetic [3, 5, 28–30, 79, 81–83] metals. Three recent candidates for plasmonic materials in the infrared—aluminum-doped ZnO (AZO), and silicon-doped InAs—are included using Drude models of recent experimental data from Naik et. al. [29], Law et al. [30], and Sachet et al. [83], respectively. For the conventional metals, data from Palik [79] was used; high-quality silver, consistent instead with the data from Johnson and Christy [85] and Wu et al. [86], would have smaller losses and a factor of three improvement in  $|\chi|^2/\text{Im}\chi$ . A broadband version of the metric can be computed for extinction or LDOS by evaluation at a single complex frequency [16, 56].

The material enhancement factor  $|\chi|^2/\text{Im}\chi$  appears in the absorption cross-section of quasispherical ellipsoids [32]; here we have shown that it more generally bounds the scattering response of a metal of any shape and size. It arises in the increased amplitude of the induced polarization currents; for example, the optimal scattering fields of Eq. (20) simplify in metals to the optimal currents

$$\mathbf{P}_{\text{scat,opt}} = \frac{i}{2} \frac{|\chi|^2}{\text{Im}\chi} \epsilon_0 \mathbf{E}_{\text{inc}}, \quad (36)$$

with similar expressions for the optimal currents for maximum absorption and LDOS. The factor  $|\chi|^2/\text{Im}\chi$  provides a balance between absorption and scattering: in terms of the polarization currents, the absorption in a metal is proportional to  $(\text{Im}\chi/|\chi|^2) \int_V |\mathbf{P}|^2$ , whereas the extinction is proportional to  $\text{Im} \int_V \mathbf{E}_{\text{inc}}^* \cdot \mathbf{P}$ , thus requiring currents proportional to  $|\chi|^2/\text{Im}\chi$  for absorption and extinction to have the same order of magnitude. The expression is intuitively appealing because a large  $|\chi|$  signifies the possibility to drive a large current, while a large  $\text{Im}\chi$  dissipates such a current. Our bounds suggest that epsilon-near-zero materials [87, 88], with  $|\chi| \approx 1$ , require a very small  $\text{Im}\chi$  to generate scattering or absorption as large as can be achieved with more conventional metals.

A similar, alternative understanding can be attained by considering the currents  $\mathbf{J} = d\mathbf{P}/dt = -i\omega\mathbf{P}$ . Defining the complex resistivity of the metal as  $\rho = i/\epsilon_0\omega\chi$ , the analogue of Eq. (36)

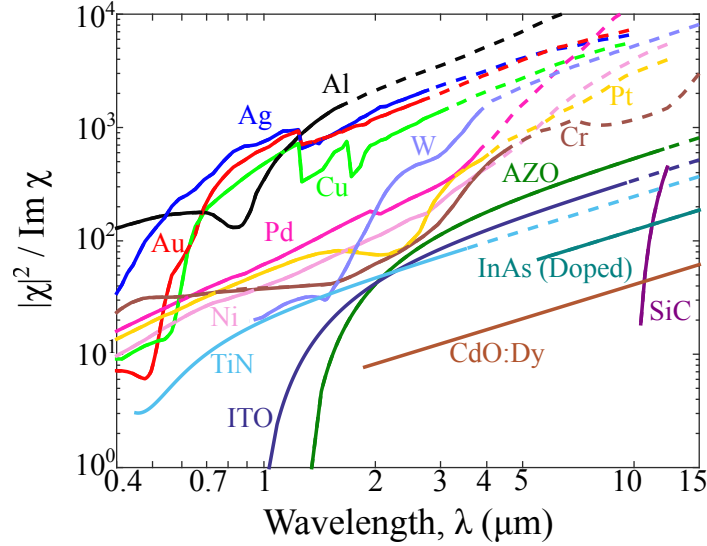


Fig. 2. A comparison of the metric  $|\chi|^2/\text{Im}\chi$ , which limits absorption, scattering, and spontaneous emission rate enhancements, for conventional metals (Ag, Al, Au, etc.) [79] as well as alternative plasmonic materials including aluminum-doped ZnO (AZO) [29], highly doped InAs [30], SiC [79], TiN [81], ITO [82], and Dysprosium-doped cadmium oxide [83] (CdO:Dy). Silver, aluminum, and gold are the best materials at visible and near-infrared wavelengths, although at higher wavelengths the structural aspect ratios needed to achieve the limiting enhancements may not be possible. The dotted lines indicate wavelengths at which resonant nanorods would require aspect ratios greater than 30, approximating the highest feasible experimental aspect ratios [84]. Despite having lower maximum enhancements, AZO, doped InAs, and SiC should be able to approach optimal enhancements in the infrared with realistic aspect ratios.

for  $\mathbf{J}$  is

$$\mathbf{J}_{\text{scat,opt}} = \frac{1}{2\text{Re}\rho} \mathbf{E}_{\text{inc}}. \quad (37)$$

The enhancement factor  $|\chi|^2/\text{Im}\chi$  thus corresponds to the inverse of the real part of the metal resistivity,

$$\frac{|\chi(\omega)|^2}{\text{Im}\chi(\omega)} = \frac{1}{\epsilon_0\omega\text{Re}\rho(\omega)}, \quad (38)$$

which corroborates the idea that small metal resistivities enable large field enhancements, as discussed recently for circuit [41] and metamaterial [6] models of single-mode response.

The limits in Eqs. (32a,32b,35a,35b) can be applied additively to multiple bodies: the per-volume absorption and scattering limits of Eqs. (32a,32b) are equally valid for a single particle, multiple closely spaced particles, layered films, or any other arrangement. Similarly, the LDOS limits in Eqs. (35a,35b) can be extended to e.g. a structure confined within two half-spaces, but with a prefactor of  $2 \times 1/8 = 1/4$ , and any other arrangement in space is similarly possible.

There are two asymptotic limits in which our bounds diverge: lossless metals ( $\text{Im}\chi \rightarrow 0$ ), and, for the LDOS bound, the limit as the emitter–metal separation distance  $d \rightarrow 0$ . In each case, the divergence is required, as there are structures that exhibit arbitrarily large responses. For

example, as the loss rate of a small metal particle goes to zero, it is known that the absorption per unit volume increases until, for a given size, the radiative loss rate equals the absorptive loss rate [43, 44, 89, 90]. However, if the size (and therefore the radiation loss rate) is decreased concurrently with the material loss, the cross-section per unit volume can be made arbitrarily large. Thus, for a small enough particle, any  $\sigma_{\text{ext}}/V$  is possible, and the limit must diverge as material loss approaches zero (regularized physically by both nonzero loss and nonlocal polarization effects [39–41]). Similarly, the LDOS can diverge in the limit of zero emitter–scatterer separation, both for lossy materials where absorption diverges [91] and for lossless materials with sharp corners [92]. The latter case can be reasoned as follows: the fields at a sharp tip, either dielectric or metal, diverge for any nonzero source (of compatible polarization) [38, 93, 94]. By reciprocity, for a source infinitesimally close to the tip, the LDOS must diverge.

#### 4. Optimal and non-optimal structures

We turn now to the design problem: are there structures that approach the limiting responses set forth by Eqs. (32a,32b,35a,35b)? We show that optimal ellipsoids can approach both the absorption and scattering limits across many frequencies by tuning their aspect ratios. For the LDOS limit, however, the optimal designs are not as clear. At the resonant (“surface-plasmon”) frequency  $\omega_{\text{sp}}$  of a given material the prototypical planar surface exhibits a nonradiative LDOS enhancement approaching the limit of Eq. (35b). However, at lower frequencies, neither thin films [95] nor common metamaterial approaches for tuning the resonant frequency achieve the  $|\chi|^2/\text{Im}\chi$  enhancement, thereby falling short of the limit. Similarly, representative designs for increased radiative LDOS are shown to fall orders of magnitude short of the limits. These structures fall short because the near-field source excites higher-order, non-optimal “dark” modes that reduce the LDOS enhancement.

To compute the electromagnetic response of the structures in Fig. 3–5, we employed a free-software implementation [96,97] of the boundary element method (BEM) [98]. Where possible (quasistatic ellipsoid extinction [32], planar metal LDOS [99]), we also used exact analytical and semi-analytical results.

##### 4.1. Absorption and scattering

Small ellipsoids approximated by their dipolar response can approach the limits of Sec. 3 across a wide frequency range by tuning their aspect ratios. Ellipsoids reach the absorption limits for small (ideally quasistatic) structures, and the scattering limits for larger structures that are still dominated by their electric dipole moment. The quasistatic absorption cross-section of an ellipsoid, for a plane wave polarized along one of the ellipsoid’s axes, is [32]

$$\frac{\sigma_{\text{abs}}}{V} = k \text{Im} \left( \frac{\chi(\omega)}{1 + L\chi(\omega)} \right), \quad (39)$$

where  $L$  is the “depolarization factor” (a complicated function of the aspect ratio) [32] along the axis of plane wave polarization. The optimal response is achieved for the aspect ratio such that  $L = \text{Re}(-1/\chi(\omega))$ , which yields a polarization field within the particle of [32]:

$$\mathbf{P} = i \frac{|\chi|^2}{\text{Im}\chi} \mathbf{E}_{\text{inc}}, \quad (40)$$

which is exactly the optimal absorption condition of Eq. (22), as can be seen by comparison with Eq. (36). For this optimal depolarization factor, the peak absorption cross-section per unit

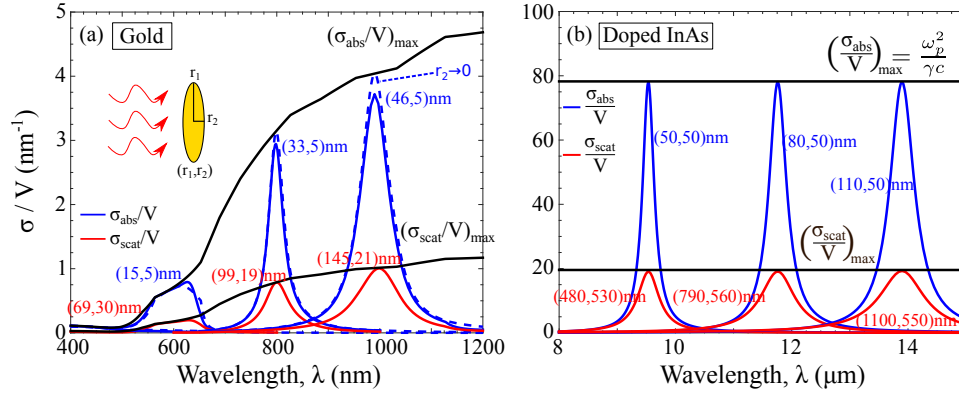


Fig. 3. Absorption (blue) and scattering (red) cross-sections per unit particle volume for nanoparticles of (a) gold [79] and (b) Si-doped InAs [30], illuminated by plane waves polarized along the particle rotation axis. The ellipsoid aspect ratios can be tuned to approach both the maximum absorption and maximum scattering cross-sections (black) of Eqs. (32a,32b). The dimensions of the nanoparticles are optimized at three representative wavelengths, in the visible for gold (constrained to have radii not less than 5nm) and at longer infrared wavelengths for doped InAs. Whereas the maximum-absorption particles are small to exhibit quasistatic behavior, represented by dashed lines in (a), the maximum-scattering particles are larger such that their scattering and absorption rates are equal.

volume is given by

$$\left[ \frac{\sigma_{\text{abs}}(\omega)}{V} \right]_{\text{ellipsoid}} = k \frac{|\chi(\omega)|^2}{\text{Im} \chi(\omega)}, \quad (41)$$

thereby reaching the general limit given by Eq. (32b). Equation (41) is valid for both oblate (disk) and prolate (needle) ellipsoids. Here we have considered the cross-section for a single incident plane wave; if one were interested in averaging the cross-section over all plane-wave angles and polarizations (as appropriate for randomly oriented particles), then it is possible to find a bound that is tighter, by 33% for most materials, in the quasistatic regime. In that case the bounds are achieved by disks but not needles.

Whereas the absorption cross-section is maximized for very small particles approaching the quasistatic limit—necessary to exhibit zero scattered power, a prerequisite for reaching the absorption bounds—the optimal scattering cross-section is achieved for larger, non-quasistatic particles that couple equally to radiation and absorption channels. One can show either through a modified long-wavelength approximation [100–102] or by coupled-mode theory [43, 44] that the dimensions of a small particle can be tuned such that the absorption and scattering cross-sections are equal, at which point the scattering cross-section per volume is a factor of four smaller than the optimal quasistatic absorption. We validate this result with computational optimization and show that it enables the design of metallic nanorods with nearly optimal performance.

Figure 3 shows the per-volume absorption and scattering cross-sections of (a) gold and (b) Si-doped InAs [30] nanorods designed for maximum response across tunable frequencies. As in Ref. [30] we employed a Drude model for the doped InAs, with plasma frequency  $\omega_p = 2\pi c/5.5\mu\text{m}$  and damping coefficient  $\gamma \approx 0.058\omega_p$ , as is appropriate for a doping density on the order of  $10^{20}\text{cm}^{-3}$  (Ref. [103]). We employed a free-software implementation [104] of the controlled random search [105, 106] optimization algorithm to find globally optimal el-

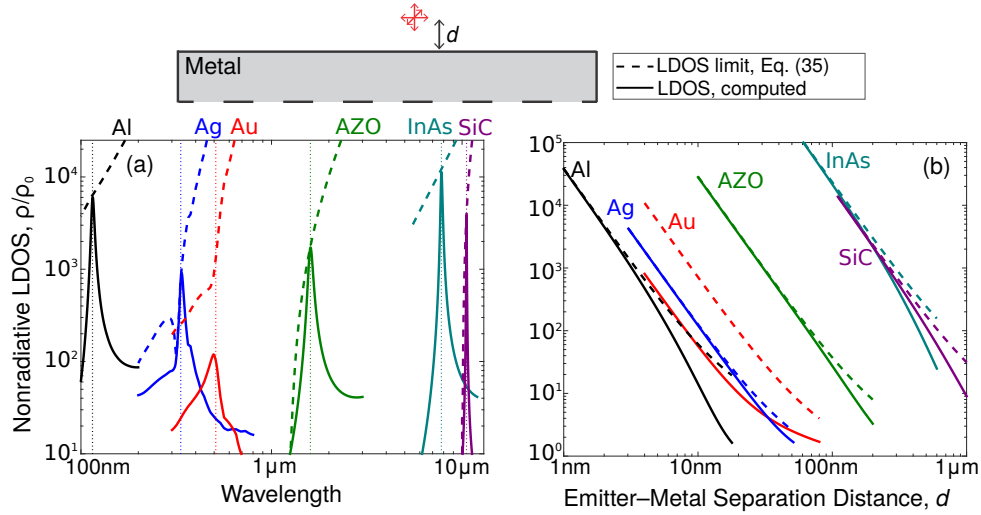


Fig. 4. Nonradiative LDOS enhancement for randomly oriented dipoles above a flat bulk metal. (a) Enhancement as a function of wavelength. For each metal except gold—which has significant losses—the nonradiative LDOS at the surface-plasmon frequency  $\omega_{\text{sp}}$  (dotted line) approaches the limit given by Eq. (35b). The emitter–metal separation distance is fixed at  $d = 0.1c/\omega_{\text{sp}}$ . The limits are equally attainable for conventional metals such as Al and Ag as for synthetic metals such as AZO [29] and highly doped InAs [30], and for SiC. (b) Enhancement as a function of metal–emitter separation distance  $d$ , with the frequency fixed at the surface-plasmon frequency  $\omega_{\text{sp}}$  for each metal. The limiting enhancements are asymptotically approached as the separation distance is decreased, because the quasistatic approximation of Eq. (42) becomes increasingly accurate.

lipsoid radii. For the gold nanorods a minimum radius of 5nm was imposed as representative of experimental feasibility [39, 107] and a size scale at which nonlocal effects are expected to remain small [39–41, 108]). The gold particles optimized for absorption fall slightly short of the limits due to the minimum-radius constraint; in the quasistatic limit (dashed), the absorption cross-section reaches the limit, as expected from Eq. (41). Both gold and doped-InAs nanoparticles closely approach the scattering and absorption limits of Eqs. (32a,32b). For Drude models, the factor  $k|\chi|^2/\text{Im}\chi$  that appears in both limits simplifies to  $\omega_p^2/\gamma c$ , a material constant independent of wavelength, clearly seen in Fig. 3(b). The increase in  $\sigma/V$  as a function of wavelength in Fig. 3(a), for gold, can be seen as a measure of the deviation of the material response [79] from a Drude model. A constant response for Drude models is not universal: *near-field* interactions, in particular the local density of states (LDOS), instead depend only on  $|\chi|^2/\text{Im}\chi \sim \omega_p^2/\gamma\omega$ , thereby increasing at longer wavelengths, away from the bulk and flat-surface plasmon frequencies.

Not all small particles reach the limiting cross-sections. Coated spheres are common structures for photothermal applications [31, 109, 110], but their absorption cross-section per unit particle volume is proportional to  $(2/3)|\chi|/\text{Im}\chi$  instead of  $|\chi|^2/\text{Im}\chi$  (cf. App. D). Their enhancement does not scale proportional to  $|\chi|^2$  due to the small metal volume fractions required to tune the resonant frequency.

#### 4.2. LDOS

Designing optimal structures for the local density of states (LDOS) enhancement limits, Eqs. (34b,34a), is not as straightforward as designing optimal particles for plane-wave absorp-

tion or scattering. Because the structure is typically in the near field of the emitter, it is difficult to design a resonant mode that exactly matches the rapidly varying field profile of the emitter. We show that it is possible to reach the nonradiative LDOS limits at the surface-plasmon frequency of a given metal, but that away from these frequencies typical structures fall short. Similarly, for the radiative LDOS limits, common structures fall short of the limits, especially at longer wavelengths.

Planar layered structures support bound surface plasmons that do not couple to radiation, and thus only improve the nonradiative LDOS. As discussed in Sec. 1, this is potentially useful for radiative heat transfer applications, where near-field emission and absorption have been extensively studied [111–115]. Moreover, adding either periodic gratings or even random textures can couple the bound modes to the far field [116–118] and potentially result in substantial increases to the radiative LDOS. Thus we first study how closely surface modes in planar structures can approach the nonradiative limits, and then we analyze the performance of representative cone- and cylindrical-antenna structures relative to the radiative LDOS limits.

At the surface-plasmon frequency, the prototypical metal-semiconductor interface that supports a surface plasmon exhibits a nonradiative LDOS approaching the limiting value of Eq. (34b). In the small-separation limit ( $kd \ll 1$ ) and at the surface-plasmon frequency  $\omega_{\text{sp}}$ , the local density of states near a planar metal interface reduces to [12]

$$\left[ \frac{\rho_{\text{nr}}(\omega_{\text{sp}})}{\rho_0} \right]_{\text{planar}} \approx \frac{1}{8(kd)^3} \frac{|\chi(\omega_{\text{sp}})|^2}{\text{Im}\chi(\omega_{\text{sp}})}, \quad (42)$$

thereby approaching exactly the nonradiative LDOS limit of Eq. (34b). (Note that we define  $\rho_0$  as the electric-only free-space LDOS, different by a factor of two from the electric+magnetic LDOS in [12].) Although the surface-plasmon frequency is typically defined [2] as the frequency at which  $\text{Re}\epsilon = -1$ , this is the frequency of optimal response only in the zero-loss limit. More generally, we define the surface-plasmon frequency  $\omega_{\text{sp}}$  such that  $\text{Re}\xi(\omega_{\text{sp}}) = \text{Re}(-1/\chi(\omega_{\text{sp}})) = 1/2$ . For gold, which never satisfies  $\text{Re}\xi = 1/2$  due to its high losses, we define surface-plasmon wavelength to be  $\lambda_{\text{sp}} = 510\text{nm}$ , where  $\text{Re}\xi(\omega)$  is a maximum.

Figure 4(a) compares semianalytical computations of the nonradiative LDOS near a flat, planar metallic interface to the nonradiative LDOS limits given by Eq. (34b). Six metals are included: Al (black), Ag (blue), Au (red), AZO (green), InAs (teal), and SiC (purple), with the surface-plasmon frequency  $\omega_{\text{sp}}$  of each in a dotted line and a fixed emitter–metal spacing of  $d = 0.1c/\omega_{\text{sp}}$ . Every metal except gold—which is too lossy—reaches its respective limit; it is possible that a different, nonplanar gold structure, with the correct “depolarization factor” (VIE eigenvalue, cf. App. A), could approach the limit. Figure 4(b) shows the emitter–metal separation–distance dependence for  $\omega = \omega_{\text{sp}}$ . The limits are approached—again, except for gold—as the emitter–metal separation decreases and the approximate  $1/d^3$  dependence of Eq. (42) becomes more accurate.

There are a few common approaches to tune the resonant frequency below  $\omega_{\text{sp}}$ . A standard approach is to use a thin film [2, 119], coupling the front- and rear-surface plasmons to create lower- and higher-frequency resonances. Other approaches include highly subwavelength structuring, to create hyperbolic [120–122] or elliptical metamaterials with reduced effective susceptibilities. We show here that such structures do not exhibit the material enhancement factor,  $|\chi|^2/\text{Im}\chi$ , and thus do not approach the limit to  $\rho_{\text{nr}}$  given by Eq. (34b).

The nonradiative LDOS of a thin film can be computed by decomposing the dipole excitation into plane waves (including evanescent waves), which reflect from the layers according to the usual Fresnel coefficients. The LDOS near a thin film is well-known as an integral over the surface-parallel wavevector [69]; for a dipole with fixed frequency  $\omega$  and height  $d$  above the



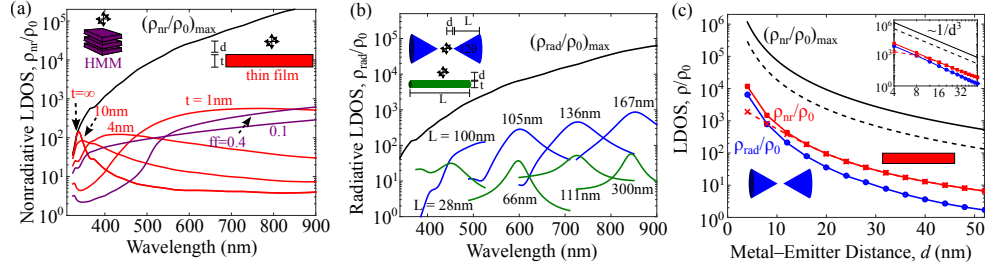


Fig. 5. Away from the surface-plasmon frequency of a given metal—taken here to be silver [79]—it is more difficult to reach the radiative and nonradiative LDOS limits, Eqs. (34b,34a). The emitter–metal separation  $d$  is fixed at  $d = 10$  nm for (a) and (b). (a) Nonradiative LDOS enhancements,  $\rho_{nr}/\rho_0$ , for thin films (red) for various silver thicknesses ( $t$ ), and for (type-I) hyperbolic metamaterials (HMMs, purple) for two silver fill fractions ( $ff$ ). (b) Radiative LDOS enhancements,  $\rho_{rad}/\rho_0$  for cone and cylinder antennas, with dimensions optimized at wavelengths from  $\lambda = 450$  nm to  $\lambda = 850$  nm. (c) Scaling of  $\rho_{nr}$  and  $\rho_{rad}$  for optimized thin films and cone antennas, respectively, as a function of  $d$  (inset: log–log scale). The scaling of the optimal design appears to be  $1/d^3$ , with the structures falling short of their respective limits [the dashed line is  $(\rho_{rad}/\rho_0)_{max}$ ] because they do not exhibit a  $|\chi|^2/\text{Im}\chi$  enhancement.

film, and a film of optimal thickness  $t$ , one can show that  $\rho_{nr}$  is (cf. App. F)

$$\left[ \frac{\rho_{nr}(\omega)}{\rho_0} \right]_{\text{thin film}} \approx \frac{1}{2(kd)^3}, \quad (43)$$

which is valid for  $\omega \ll \omega_{sp}$  (otherwise a bulk is optimal and Eq. (42) describes the response) and for relatively small loss,  $\text{Im}\chi \ll |\text{Re}\chi|$  (as is typical at optical frequencies), to ensure the large-wavevector modes are resolved. Unlike a planar interface, an optimal thin film does not exhibit the  $|\chi|^2/\text{Im}\chi$  material enhancement. The thin film falls short because it relies on near-field interference to couple the front- and rear-surface plasmons, yielding a resonance that couples strongly to the dipole emitter over only a small bandwidth of wavevectors ( $\Delta k_p \sim \text{Im}\chi/|\chi|$ ) that cancels the resonant enhancement from decreased loss. An alternative understanding arises from viewing a thin film as the single-unit-cell limit of a layered hyperbolic metamaterial (HMM) [95]. Hyperbolic metamaterials exhibit anisotropic effective susceptibilities such that the resonant frequency can be tuned, but their resonances occur within the bulk rather than along the surface, such that they cannot yield infinite LDOS even in the limit of zero loss. In Ref. [95] we show that the LDOS near an optimal HMM is nearly identical to the LDOS near an optimal thin film, as verified in Fig. 5.

Achieving the radiative LDOS limit is a similarly challenging problem. To reach the radiative LDOS limits, the polarization field must exhibit the  $1/r^3$  spatial dependence of the incident field, while also coupling to far-field radiation channels. Here we consider two representative structures for tuning the LDOS resonant frequency: mirror-image cones (akin to bowtie antennas) and cylindrical antennas (scaled [66] shorter than  $\lambda/2$ ), and we show that each falls short of the limits across optical frequencies.

Figure 5 shows the nonradiative and radiative LDOS near silver thin films, effective-medium hyperbolic metamaterials (HMMs), cones, and cylinders. In Fig. 5(a) and Fig. 5(b), the emitter–metal separation distance is fixed at  $d = 10$  nm, and the structures are optimized at four wavelengths,  $\lambda = [450, 600, 725, 850]$  nm, using a standard local optimization algorithm [123]. The optimal lengths and thicknesses are included in the figure. The optimal cone half-angles and

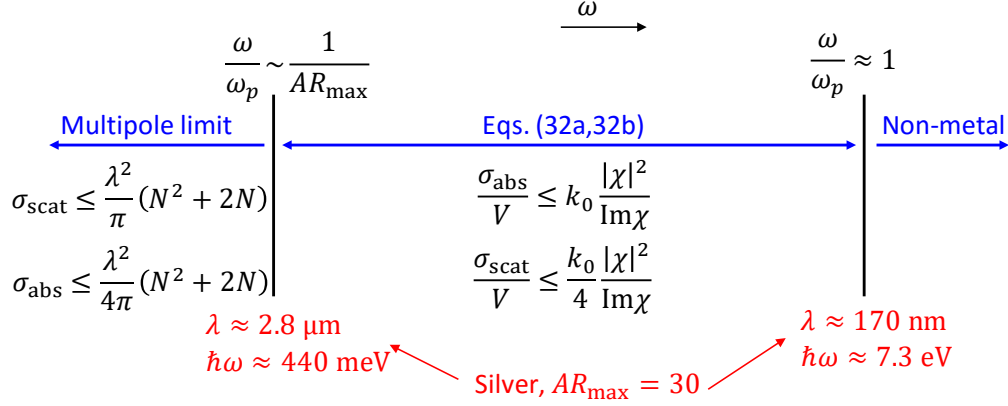


Fig. 6. A schematic comparison of absorption and scattering limits: multipole limits [21, 22] to the total cross-section can provide design guidelines at low frequencies, where it is difficult to achieve “plasmonic” resonances, but at higher frequencies our dissipation-based limits provide tighter limits to the per-volume response. The frequencies at which our bounds can be reached range from the bulk plasma frequency ( $\omega = \omega_p$ ) down to  $\omega \sim \omega_p/AR_{\max}$ , where  $AR_{\max}$  is the maximum achievable aspect ratio. Included is the relevant range for silver ellipsoids (red text), assuming  $AR_{\max} = 30$ . Plasmonic behavior at longer wavelengths is possible with materials such as AZO and doped InAs (cf. Fig. 2).

cylinder radii vary slightly but are typically on the order of  $\theta = 20^\circ$  and  $r = 10\text{nm}$ , resp. Away from  $\omega_{\text{sp}}$ , the optimal nonradiative LDOS of the thin films in Fig. 5(a) is very well described by Eq. (43), which shows that the structures fall short of the limits by the material enhancement factor. A similar effect is seen for the radiative LDOS in Fig. 5(b), as the cones also fall increasingly short of the limits as the frequency is decreased ( $\approx 250\times$  at  $\lambda = 850\text{nm}$ ). The quantum efficiencies ( $\rho_{\text{rad}}/\rho_{\text{tot}}$ ) of the cone antenna varies from 50–70%, near the optimal ratio of 50% to reach our bounds, suggesting that the reason they fall short is due to a mismatch between the emitter and resonance field profiles, not due to coupling to nonradiative channels. Figure 5(c) shows the LDOS dependencies as a function of  $d$  for  $\lambda = 600\text{nm}$ , in linear (main) and logarithmic (inset) scale, with the structure parameters optimized for each  $d$ . The optimal structures appear to exhibit the  $1/d^3$  scaling of the limits in Eqs. (34a,34b). An important open question is the extent to which structures can be designed to approach the limits, thereby improving over current designs by two to three orders of magnitude.

## 5. Extensions and discussion

The limits derived in Secs. (2,3) apply to a general class of materials embedded in vacuum, without any scatterers. They can be generalized for non-vacuum backgrounds. For a metal in a homogeneous, lossless background with permittivity  $\epsilon_{\text{bg}}$ , the limits are of the same form but with replacements  $k \rightarrow \sqrt{\epsilon_{\text{bg}}}k$  and  $\chi \rightarrow (\epsilon(\omega) - \epsilon_{\text{bg}})/\epsilon_{\text{bg}}$ . A similar generalization applies for general lossy media in non-vacuum backgrounds. If there are background scatterers present (possibly periodic [124–126]), then the derivation is identical except that the Green’s functions are the Green’s functions in the presence of the background scatterers, and the “incident” fields therefore incorporate the effects of the background scatterers. The minimum thickness of a metal absorber on a substrate [127, 128], for example, could be computed with Eq. (32b), with the replacement  $E_{\text{inc}} = E_0 (e^{ikz} + r e^{-ikz})$  (for substrate reflectivity  $r$ ), or more generally by bounding the incident field via  $|E_{\text{inc}}| < 2|E_0|$ .

Previous approaches to general limits via energy conservation have bounded the response of a structure via its scattered-field operator [27]. This has yielded limits to the cross-section,  $\sigma$ , for spherically symmetric [43, 44] or more general [21, 22] scatterers whose response has been decomposed into spherical multipoles. The cross-section limits are proportional to  $\lambda^2 (N^2 + N)$ , where  $N$  is the number of excited multipoles.

From a design perspective, there are serious impediments to using such cross-section limits (as opposed to our  $\sigma/V$  limits). The cross-section itself is unbounded, increasing with the size of a large particle [32]. Thus implicit in any use of such bounds is a size normalization, but this can only be done at very small scales, where the number of multipoles is 1 (or 2 for perfect conductors) and our bounds are tighter, and at very large size scales, approaching the geometric-optics regime. At intermediate sizes, it is difficult to estimate the number of multipoles without further modeling of a given structure. We have presented a new approach, using material dissipation, instead of the scattering operator, as the binding constraint. Our limits have the unique feature that they incorporate the material properties and are independent of structure. From a design viewpoint this is a significant advantage, since for any problem there are infinitely many possible structures but typically only a handful of relevant materials. Furthermore, the normalization to geometric volume, e.g.  $\sigma/V$ , emerges naturally in our limits.

Reaching our absorption and scattering limits likely requires significant polarization currents throughout the scatterer volume, as can be seen in the optimal field profiles in Sec. 3. At low frequencies, it is difficult to fabricate structures with sizes or aspect ratios necessary to achieve such resonances. For a Drude-metal (appropriate at low frequencies) ellipsoidal nanorod, the optimal aspect ratio [32] for maximum absorption scales as  $\omega/\omega_p$ , where  $\omega_p$  is the bulk plasma frequency. Thus, if we define a maximum feasible aspect ratio  $AR_{\max}$ , the minimum frequency at which a plasmonic resonance can be achieved is proportional to  $\omega_p/AR_{\max}$ . Below this frequency, the multipole limits can serve as a design guide, although the uncertainty about the number of multipoles, and the potential mismatch between a non-spherical object and its bounding sphere, remain as barriers. Figure 6 depicts schematically which bound provides better design criteria as a function of frequency. Included in Fig. 6 is red text corresponding to the limiting frequencies at which silver nanorods (using experimental material data [79] and making no Drude approximation) can approach our limits, assuming a realistic [84] maximum aspect ratio of 30.

Similarly, it may be difficult to reach our limits with larger, wavelength-scale solid particles that are much larger than the skin depth. One of the conclusions from our work is that such particles are particularly inefficient scatterers, and thus should be *avoided*, because currents cannot be excited throughout such a large portion of their volume. At optical frequencies, any technology must ultimately incorporate some *collection* of ordered or disordered scatterers, whether in planar arrays [4, 57, 129–132], aqueous environments [133–135], or some other configuration. Thus even if the individual scatterers have small cross-sections, there can be many of them (due to their small volumes), providing a large collective cross-section [107] while maintaining the per-volume response of the individual scatterers.

Aside from particle scattering, our limits extend to situations that do not have multipole counterparts. They yield meaningful limits for extended structures (whose large size would excite many spherical harmonics), and for the LDOS (where a near-field source would excite many spherical harmonics).

An interesting aspect of our limits is the fact that they apply to any open scattering problem. Open systems—in which energy can enter and exit—are typically described by non-Hermitian operators. Non-Hermitian operators are not guaranteed to be diagonalizable, and therefore may not have a complete basis of eigenfunctions (technically, even for Hermitian operators, rigorous eigendecomposition in infinite-dimensional spaces is subtle and subject to obscure counter-

examples [136, 137]). Breakdown of diagonalizability only occurs at “exceptional points” that must be forced [138, 139] and which occur by chance with zero probability. Near an exceptional point, however, it is possible to have eigenfunctions that are nearly “self-orthogonal,” with exceptionally large modal overlap (or in theory highly nonnormal and ill-conditioned Maxwell operators [140]) leading to effects such as destructive interference in scattering “dark states” [141] (e.g. Fano resonances [142]) and the Petermann factor for noise enhancement in lasers [143, 144]. Nevertheless, passive systems near such exceptional points cannot exceed our limits, which impose only conservation of energy.

The limits presented here suggest new design opportunities with metals. Nearly lossless metals [8] could manifest unprecedented responses. Even for conventional lossy metals, large-area structures that achieved the absorption or scattering limits presented here could potentially do so with thicknesses approaching a single atomic layer. Nonlocal interactions, for which it may not be possible to separate material and structural properties, would be important in such a structure. Finding limits incorporating nonlocal effects would represent an important extension to this work. Designing structures to approach the LDOS limits could impact applications such as imaging, where there are potentially orders of magnitude improvement to be gained. We derive limits for the problem of near-field radiative heat transfer, where the sources are embedded *within* the designable media, in an upcoming publication [145], and it would be interesting to extend the limits derived here to other figures of merit, potentially finding new metrics or structures for optimal light–matter interactions.

### A. Alternative understanding of the limits: VIE approach

Here we present an alternative viewpoint for understanding our limits, as arising from sum rules over eigenmodes of the volume integral equations of electromagnetism. This approach appears to only work for materials with a scalar  $\chi$  (either electric or magnetic), and thus is less general than the derivation in the text (we assume a nonmagnetic medium). We include this appendix because higher-order sum rules may yield tighter limits in certain scenarios, e.g. angle-averaged incident fields. First we show how limits arise from eigenmodes of the volume integral equations (VIEs), which can be considered “material resonances.” This connection was partially recognized by Rahola [146], and may be related to eigenvalues in SALT laser theory [147], but has not since been pursued any further. Material resonances are common in quasistatic electromagnetism [148, 149], where there is no frequency; here, we show how to extend the concept to fixed, nonzero frequencies.

Inherent to the concept of a resonance in physics is the resonant frequency: intuitively, the frequency at which an electromagnetic, elastic, quantum mechanical, or any other type of wave oscillates without external forcing in a specific, predefined structure. In a closed or periodic structure, these resonances correspond mathematically to eigenvalues of the underlying differential equations. For photonic structures, defined by a spatially dependent permittivity  $\epsilon(\mathbf{x})$ , resonant frequencies  $\omega_n$  are eigenvalues of the eigenequation

$$\frac{1}{\epsilon(\mathbf{x})} \nabla \times \nabla \times \mathbf{E}_n = \left( \frac{\omega_n}{c} \right)^2 \mathbf{E}_n, \quad (44)$$

defined with appropriate boundary conditions and the divergence equation  $\nabla \cdot \epsilon \mathbf{E} = 0$ . For a material with negligible dispersion (metals introduce further complications), the Maxwell operator  $\mathcal{M} = \epsilon^{-1} \nabla \times \nabla \times$  is frequency-independent, such that the eigenvalue corresponds to a resonant frequency. In open systems, resonances are complex poles of the Green’s function (or of the scattering operator) rather than true eigenvalues, because their “eigenfunctions” (“leaky modes”) diverge exponentially in space [150, 151]. For a sign convention of  $e^{-i\omega t}$  time dependence, these poles  $\omega_n$  must lie in the lower-half of the complex-frequency plane as shown

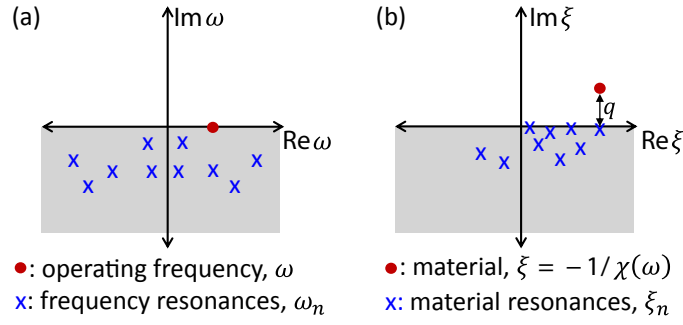


Fig. 7. A comparison of the (a) resonant-frequency and (b) resonant-susceptibility frameworks in electromagnetism. The conventional resonant-frequency approach is depicted in (a): the operating frequency is real-valued and can in theory be approached arbitrarily closely by a resonance with small imaginary part  $-\text{Im } \omega_n$  (i.e. a high-Q resonance). Conversely, volume integral equations yield the resonant-susceptibility approach depicted in (b): metal losses, which correspond to  $\text{Im } \xi = \text{Im}(-1/\chi) > 0$ , inherently impose a minimum separation  $q$  to how closely a material resonance—restricted to lie on or below the real line—can approach the real system parameters. Moreover, quasistatic structures have real-valued eigenvalues, and thus have the potential to achieve the minimum eigenvalue separation and maximum optical response.

in Fig. 7(a). Frequency resonances in electromagnetism are well understood, but *material* resonances—which arise as eigenvalues in integral equations—have hardly been explored at all.

The electric field integral equation (EFIE) formulation of Maxwell’s equations is derived through the use of Green’s functions [51, 152]. As depicted in Fig. 1, we consider generic scattering problems in which a structure with susceptibility  $\chi$  interacts with an externally imposed incident field  $\mathbf{E}_{\text{inc}}$ . The response of the scatterer is given by a convolution of the free-space Green’s function  $\bar{G}$  with the induced polarization currents  $\mathbf{P} = \chi\mathbf{E}$ . As in the text we assume a scatterer embedded in vacuum, with straightforward generalizations. The total field  $\mathbf{E}$  is the sum of the incident and scattered fields [51, 153],

$$\mathbf{E}(\mathbf{x}) = \mathbf{E}_{\text{inc}}(\mathbf{x}) - \int_V \chi(\mathbf{x}') \bar{G}(\mathbf{x}, \mathbf{x}', \omega) \mathbf{E}(\mathbf{x}') dV \quad (45)$$

for all points in space, where we choose a negative sign convention for the Green’s function (opposite that of Eq. (10) in the text). Equation (45) can be desingularized [154], but our treatment depends only on an abstract spectral decomposition that does not require us to grapple with such details. A similar integral equation arises in quantum mechanics, where it is known as the Lippmann–Schwinger equation and the susceptibility is replaced by the scattering potential [155]. For a scatterer with homogeneous susceptibility,  $\chi$  is constant and can be taken out of the integrand in Eq. (45).

Homogeneous scatterers are thus defined by the single material parameter  $\chi$ . Resonances of Eq. (45), even in open systems, are true eigenvalues because the integral equation has unknowns  $\mathbf{E}(\mathbf{x}')$  defined over the finite scatterer domain  $V$ , and thus the corresponding eigenfunctions are normalizable. Eigenfunctions  $\mathbf{E}_n$  and eigenvalues  $\xi_n$  of the Green’s function integral operator satisfy:

$$\int_V \bar{G}(\mathbf{x}, \mathbf{x}', \omega) \mathbf{E}_n(\mathbf{x}') dV = \xi_n \mathbf{E}_n(\mathbf{x}) = -\frac{1}{\chi_n} \mathbf{E}_n(\mathbf{x}) \quad (46)$$

for all points  $\mathbf{x}$  in  $V$ . Given the eigenvalue  $\xi_n$ , if we choose a material  $\chi = \chi_n = -1/\xi_n$  then by comparison with Eq. (45) we see that a “standing-wave”  $\mathbf{E} \neq 0$  is possible even for  $\mathbf{E}_{\text{inc}} = 0$ . Equation (46) is the integral-equation analogue of Eq. (44) (for a homogeneous scatterer), and yet we see that the integral operator on the left-hand side of Eq. (46) depends on the structure and the frequency but not on the susceptibility. Instead, the eigenvalue of the mode yields a resonant value  $\chi_n$ —a resonant material susceptibility, for a fixed frequency. Just as “leaky modes” in Eq. (44) are not actually physical solutions of Maxwell’s equations, the resonances  $\chi_n$  are not physically realizable materials; we will see below that  $\text{Im } \chi_n < 0$  for  $\omega > 0$ , as shown in Fig. 7(b), corresponding to gain required to overcome modal radiation loss.

For both frequency and material resonances, the separation between the system parameter (e.g. operational frequency) and the resonance defines the magnitude of the response—the smaller the separation, the larger the response. In the resonant-frequency framework it is difficult to provide a lower bound on the imaginary part of the resonant frequency  $\omega_n$  (thereby bounding the maximum  $Q$ ). In the resonant-material approach, however, the system parameter—the susceptibility, instead of the frequency—has a nonzero imaginary part for a lossy system. By causality [156] (or passivity [52]), frequency resonances for a fixed structure and material lie in the lower half of the complex- $\omega$  plane. Similarly, for a fixed frequency  $\omega > 0$ , the material resonances  $\xi_n$  must reside in the lower half of the complex- $\xi$  plane (otherwise, one could construct a passive linear material that violates the condition on the resonant frequencies [157]). Thus, as depicted in Fig. 7, there is a minimum separation  $q = \text{Im } \xi(\omega)$  between the material parameter  $\xi(\omega) = -1/\chi(\omega)$  and the eigenvalue  $\xi_n$ . There are further benefits to the resonant-material approach: the solutions to the integral equation are defined only over the scatterers, rather than all space, and quantities like the extinguished power and the local density of states can be written as volume integrals, ideally suited to a VIE framework.

To simplify further analysis, we rewrite the VIE of Eq. (45) as:

$$(\mathcal{I} + \chi\mathcal{G})\mathbf{e} = \mathbf{e}_{\text{inc}} \quad (47)$$

where the fields are  $\mathbf{e}$  and  $\mathbf{e}_{\text{inc}}$  (with lower-case  $\mathbf{e}$  denoting vector fields restricted to the volume  $V$ , forming a Hilbert space, as opposed to the fields  $\mathbf{E}$  defined everywhere in space),  $\mathcal{I}$  is the identity operator, and  $\mathcal{G}$  is the Green’s function integral operator defined by  $\mathcal{G}\mathbf{e} = \int_V \overline{\overline{G}}(\mathbf{x}, \mathbf{x}', \omega)\mathbf{E}(\mathbf{x}')$ . The eigenfunctions of  $\mathcal{G}$  are solutions of Eq. (46) and are given in vector notation by

$$\mathcal{G}\mathbf{e}_n = \xi_n\mathbf{e}_n, \quad (48)$$

where  $n$  is the mode index,  $\xi_n = -1/\chi_n$ , and  $\text{Im } \xi_n \leq 0$ . There are advantages to considering the integral operator  $\mathcal{G}$  rather than the Maxwell operator  $\mathcal{M} = (1/\epsilon)\nabla \times \nabla \times$ . For the operator  $\mathcal{M}$ , metals are difficult to treat: material dispersion renders the eigenvalue problem nonlinear in  $\omega^2$ , and material loss yields a non-Hermitian operator even for closed or periodic structures [158–160]. The Green’s function operator avoids these difficulties because it assumes a fixed frequency and is independent of material, and therefore of material loss. An additional advantage of the VIE approach is the compact domain, which sidesteps the subtle normalization [159, 161–164] required in the resonant-frequency approach, where the “leaky” fields diverge as they extend to infinity.

We consider open systems—in which energy can enter and exit—such that typical operators, including the curl-curl operator  $\mathcal{M}$  of Eq. (44) and the Green’s function integral operator  $\mathcal{G}$  of Eq. (47), are not Hermitian. By reciprocity  $\mathcal{G}$  is complex-symmetric [11], such that if it is diagonalizable, its generic  $\mathcal{U}\Xi\mathcal{U}^{-1}$  eigendecomposition can be written [165]

$$\mathcal{G} = \mathcal{U}\Xi\mathcal{U}^T, \quad (49)$$

where  $\Xi$  is a diagonal operator with entries  $\xi_n$  and  $\mathcal{U}$  is the basis of eigenfunctions  $\mathbf{e}_n$ . The assumption of diagonalizability (the existence of a “spectral” eigendecomposition of the operator) is commonplace in physics. Technically, even for Hermitian operators, rigorous eigendecomposition in infinite-dimensional spaces is subtle and subject to obscure counterexamples [136, 137]. However, if one makes the reasonable conjecture that the system can be simulated on a computer, i.e. that one has a convergent finite-dimensional discretization, then one should be able to apply the eigendecomposition in this finite-dimensional system to arbitrary accuracy; this can be viewed as a justification of the commonplace assumption of diagonalizability. Even in a finite-dimensional problem, of course, diagonalizability of a non-Hermitian matrix is not guaranteed [165], but breakdown of diagonalizability only occurs at “exceptional points” that must be forced [138, 139]. Exceptional points occur by chance with zero probability; near an exceptional point, the operator  $\mathcal{G}$  is highly nonnormal [140] and ill-conditioned (i.e.  $\text{cond}(\mathcal{U}\mathcal{U}^\dagger) \gg 1$ ), but it is diagonalizable, and we show in the text that the response of such structures cannot surpass our limits. By continuity our limits also apply at non-diagonalizable exceptional points, thereby justifying our assumption of diagonalizability. Our decomposition is similar in spirit to the singularity and eigenmode expansion methods, SEM and EEM, respectively, for surface integral equations in electromagnetic scattering [68, 166–169].

The complex-symmetry of the operator leads to an atypical (unconjugated, indefinite) “inner product” under which the eigenfunctions are orthogonal:  $\mathbf{e}_i^T \mathbf{e}_j = \int_V \mathbf{E}_i \cdot \mathbf{E}_j = \delta_{ij}$ . Because modes are not orthogonal under the typical inner product  $\mathbf{e}^\dagger \mathbf{e} = \int_V \mathbf{E}^* \cdot \mathbf{E}$ , they are not *power*-orthogonal [162]: it is possible for energy in one mode to mix into another, leading to effects such as destructive interference in scattering “dark states” [141] (e.g. Fano resonances [142]) and the Petermann factor for noise enhancement in lasers [143, 144]. Another possibility is that  $\int_V \mathbf{E}_i \cdot \mathbf{E}_i = 0$ , i.e. “self-orthogonality,” which renders the eigenfunction basis incomplete [170] and corresponds to exceptional points [139, 171]. As shown in the text, energy conservation prevents any of these phenomena from surpassing our limits.

A modal decomposition of the electric field  $\mathbf{e}$  follows from the modal decomposition of  $\mathcal{G}$  in Eq. (49). In order to isolate  $\mathcal{G}$ , it is easier to work with the polarization currents,  $\chi \mathbf{e}$ , which from Eq. (47) are given by:

$$\begin{aligned} \chi \mathbf{e} &= \chi(\omega) (\mathcal{S} + \chi(\omega)\mathcal{G})^{-1} \mathbf{e}_{\text{inc}} \\ &= \mathcal{U} (\Xi - \xi(\omega))^{-1} \mathcal{U}^T \mathbf{e}_{\text{inc}}, \end{aligned} \quad (50)$$

where  $\xi(\omega) = -1/\chi(\omega)$  and the inverse of  $\mathcal{S} + \chi\mathcal{G}$  is guaranteed to exist (for  $\varepsilon \neq 0, 1$ ) because it is a Fredholm operator [172].

Because the extinguished power and the total LDOS are both imaginary parts of linear functionals of the induced fields, their modal decompositions are single sums over the resonances. Given the electric field from Eq. (50), the extinction of Eq. (5) and total LDOS of Eq. (15) are determined by overlap integrals between the incident field and the VIE-basis eigenfunctions:

$$P_{\text{ext}} = \frac{\varepsilon_0 \omega}{2} \text{Im} \left[ \mathbf{e}_{\text{inc}}^\dagger \mathcal{U} (\Xi - \xi(\omega))^{-1} \mathcal{U}^T \mathbf{e}_{\text{inc}} \right] \quad (51a)$$

$$\frac{\rho_{\text{tot}}}{\rho_0} = 1 + \frac{2\pi n_0 \omega}{c} \text{Im} \left[ \mathbf{e}_{\text{inc}}^T \mathcal{U} (\Xi - \xi(\omega))^{-1} \mathcal{U}^T \mathbf{e}_{\text{inc}} \right]. \quad (51b)$$

Because  $(\Xi - \xi(\omega))^{-1}$  is a diagonal operator, it can be written as a single sum over the modes,

simplifying the extinction and LDOS:

$$P_{\text{ext}} = \frac{\epsilon_0 \omega}{2} \text{Im} \sum_n \frac{p_n}{\xi_n - \xi(\omega)} \quad (52a)$$

$$\frac{\rho_{\text{tot}}}{\rho_0} = 1 + \frac{2\pi}{k^3} \text{Im} \sum_n \frac{\rho_n}{\xi_n - \xi(\omega)} \quad (52b)$$

where  $p_n = \int_V \mathbf{E}_{\text{inc}}^* \cdot \mathbf{E}_n \int_V \mathbf{E}_n \cdot \mathbf{E}_{\text{inc}}$ ,  $\rho_n = \sum_j \left( \int_V \mathbf{E}_n \cdot \mathbf{E}_{\text{inc},s_j} \right)^2$ , and  $\xi_n$  are the VIE eigenvalues and the diagonal entries of  $\Xi$ . The  $p_n$  and  $\rho_n$  are normalized ‘‘oscillator strengths’’ representing the per-mode extinguished power and per-mode total density of states, respectively. Adding up only the oscillator strengths yields structure-independent sum rules for each quantity:

$$\sum_n p_n = \int_V |\mathbf{E}_{\text{inc}}|^2 dV \quad (53a)$$

$$\sum_n \rho_n = \sum_j \int_V \mathbf{E}_{\text{inc},s_j} \cdot \mathbf{E}_{\text{inc},s_j} dV, \quad (53b)$$

where the sum over all modes exploits the eigenbasis orthogonality,  $\mathcal{U}^T \mathcal{U} = \mathcal{I}$ , and as before  $j$  indexes the polarization of the dipole emitter. For each quantity, the sum of the oscillator strengths  $p_n$  or  $\rho_n$  is given by the intensity of the electric field originally incident upon the volume occupied by the scatterer.

Ideally, the sum rules for the extinction and the total LDOS would lead directly to limits, with the numerators in Eqs. (52a,52b) bounded above by the sum rules and the denominators bounded below by  $\text{Im} \xi(\omega) = \text{Im} \chi(\omega) / |\chi(\omega)|^2$ , which is nonzero due to material losses. Because our system is non-Hermitian, such an argument is not valid. If we had a Hermitian system with a conjugated orthogonality relationship  $\int \mathbf{E}_i^* \cdot \mathbf{E}_j = 0$ , then we would have used  $\mathcal{U}^\dagger$  instead of  $\mathcal{U}^T$  in Eq. (49) and the resulting amplitudes  $p_n$  and  $\rho_n$  would have been real and positive. Due to radiation losses (though *not* metal losses), we have complex  $p_n$  and  $\rho_n$  with possibly negative real parts, and hence it is possible to have e.g.  $|p_n| \gg |\sum p_n|$ . Such a response is a general feature of *nonnormal* dynamics, where there can be significant amplification beyond what one would expect from the resonances, as the pseudospectral level curves [140] may look very different from typical circles centered at the eigenvalues.

Energy conservation prevents such responses from surpassing the sum rule limits that are obtained when all of the oscillator strength in Eqs. (53a,53b) are concentrated at a single resonance (with the caveat that  $\mathbf{E}_{\text{inc},s_j} \rightarrow \mathbf{E}_{\text{inc},s_j}^*$ , although for the dominant  $1/r^3$  quasistatic term this makes no difference). The optimal resonance is given by

$$\xi_{\text{ext,opt}} = \text{Re} \xi(\omega) \quad (54)$$

which is on the real line, as close to the material parameter  $\xi(\omega)$  as possible. This choice of eigenvalue leads to the extinction limit in Eq. (32b). Imposing energy conservation on the absorption, scattering, and radiative and nonradiative LDOS integrals in the VIE approach yields the limits of Eqs. (32a,32b,34a,34b), but we will not prove that here.

An interesting possibility that arises from the VIE approach is the potential existence of further sum rules. In addition to considering the sum of oscillator strengths in Eq. (53a), one



can consider the sum of eigenvalues, weighted by the oscillator strengths:

$$\begin{aligned}
& \sum_n \xi_n p_n \\
&= \sum_n \int \int \mathbf{E}_i^{\text{inc}}(\mathbf{x}) \mathbf{E}_i^n(\mathbf{x}) \xi_n \mathbf{E}_j^n(\mathbf{x}') \mathbf{E}_j^{\text{inc}}(\mathbf{x}') \\
&= \sum_n \int \int \int \mathbf{E}_i^{\text{inc}}(\mathbf{x}) \mathbf{E}_i^n(\mathbf{x}) \bar{\bar{G}}_{jk}(\mathbf{x}', \mathbf{x}'') \mathbf{E}_k^n(\mathbf{x}'') \mathbf{E}_j^{\text{inc}}(\mathbf{x}') \\
&= \int \int \mathbf{E}_i^{\text{inc}}(\mathbf{x}) \bar{\bar{G}}_{ji}(\mathbf{x}', \mathbf{x}) \mathbf{E}_j^{\text{inc}}(\mathbf{x}')
\end{aligned} \tag{55}$$

where we used the resolution of the identity  $\sum_n \mathbf{E}_i^n(\mathbf{x}) \mathbf{E}_k^n(\mathbf{x}'') = \delta_{ik} \delta(\mathbf{x} - \mathbf{x}'')$  to simplify the third line. In the surface-integral representation of quasistatic electromagnetism it can be shown the oscillator strengths and relevant eigenvalues for extinction averaged over all angles is constrained to satisfy  $\sum_n \xi_n p_n / \sum_n p_n = 1/3$ , reducing the possibility for all-angle response relative to the single-angle response [45, 47]. Equation (55) and its higher-order counterparts (e.g.  $\sum_n \xi_n^2 p_n \sim \mathbf{e}_{\text{inc}}^\dagger \mathcal{G} \mathcal{G} \mathbf{e}_{\text{inc}}$ ) may yield stricter sum rules under various incident fields, reducing the possible response.

### B. Bounds on the $O(kL)$ term in the LDOS limits

The bounds on the radiative and nonradiative LDOS, Eqs. (34a,34b), take into account the  $1/r^3$ ,  $1/r^2$ , and  $1/r$  terms in the free-space dyadic Green's function. Integrating the  $1/r$  term over a half-space (or a spherical shell, or any other structure separated some distance  $d$  from the source) yields the  $O(kL)$  term in Eqs. (34a,34b), which diverges as the size  $L \rightarrow \infty$ . As discussed in the text, this divergence is unphysical. It results from deriving the optimal current as proportional to the incident field,  $\mathbf{P} \sim \mathbf{E}_{\text{inc}}$ , which is appropriate and feasible for the evanescent waves, but which for the  $1/r$  term yields a physically impossible fixed energy density over infinite space within a lossy medium.

Despite the divergence of this term, for *finite* object sizes its contribution to the limits of Eqs. (34a,34b) is actually very small. One does not even have to consider a finite object but a finite interaction distance:  $L$  represents the largest distance over which polarization currents in the metal generate nonzero scattered fields at the dipole source. Even wavelength-scale lengths are upper bounds to reasonable interaction distances in a lossy medium. Table 1 compares the near-field limit in Eq. (35a) to the full limit with all terms in Eq. (34a) as well as the full limit without the  $O(kL)$  term,

$$\frac{\rho_{\text{nr}}}{\rho_0} \leq \frac{|\chi|^2}{\text{Im} \chi} \left[ \frac{1}{32(kd)^3} + \frac{1}{16kd} \right] + 1. \tag{56}$$

We take  $|\chi|^2 / \text{Im} \chi = 1$  for simplicity. One can see from Table 1 that the near-field limit is a very good approximation to the overall limit for realistic separations, and that the far-field term contributes at most 0.03% for the cases considered.

### C. Suppressing absorption

As discussed in the text, the optimal scattering limits are reached when absorption and scattering are exactly equal. For some applications (e.g. solar cells enhanced by plasmonic particle scatterers [4]), however, parasitic absorption is detrimental and needs to be avoided, even if the per-volume scattering is reduced. Here we present alternate limits to Eq. (32a) to account for absorption suppression.

Table 1. Tabulation of higher-order term contributions to radiative LDOS limits.

kd	kL	Eq. (35a)	Eq. (56)	Rel. Error	Eq. (34a)	Rel. Error
0.0628	1	125.98	127.97	1.56%	128.02	+0.03%
0.01	1	31250	31257.25	0.0232%	31257.3	+0.0001%
0.01	10	31250	31257.25	0.0232%	31257.68	+0.0013%

We define a fraction  $f$  that is the ratio of absorption to extinction,

$$f = \frac{P_{\text{abs}}}{P_{\text{abs}} + P_{\text{scat}}}. \quad (57)$$

Suppose we define our figure of merit as the maximum scattering cross-section per unit volume subject to the condition that  $f$  is smaller than some maximum ratio  $f_{\text{max}}$ :

$$\begin{aligned} \max \quad & P_{\text{scat}} \\ \text{s.t.} \quad & f < f_{\text{max}}. \end{aligned}$$

As in Sec. 3, we use standard Lagrangian optimization techniques and the the fact that the constraint on  $f$  is active ( $f = f_{\text{max}}$ ) by the KKT conditions. By the same steps as in Sec. 3, it is straightforward to show that the per-volume scattering cross-section is limited by:

$$\frac{\sigma_{\text{scat}}}{V} \leq f_{\text{max}} (1 - f_{\text{max}}) k \frac{|\chi(\omega)|^2}{\text{Im} \chi(\omega)}. \quad (58a)$$

Eq. (58a) is maximized at  $f_{\text{max}} = 0.5$ , corresponding to the unconstrained optimum in Eq. (29a). For significantly reduced absorption (say  $< 2\%$ ), there is a significant penalty ( $> 10\times$ ) to the maximum per-volume scattering.

#### D. Quasistatic cross-section of a coated sphere

Also shown in Fig. 3 are the cross-sections of coated spheres that have been optimized by the same procedure as the ellipsoids. Coated spheres, with dielectric cores and metallic shells, are a common structure for tunable resonances across visible and infrared frequencies [31, 110, 173–175], but one can see that their response falls short of the limits. To understand why the performance falls short, we consider the quasistatic absorption cross-section, which is known analytically [32] and can be compared to Eq. (39). For simplicity we assume the particle core has the same permittivity as the shell, which has only a small effect on the response but enables us to write the typical [32] coated-sphere cross-section  $\sigma_{\text{cs}}$  in the form

$$\left( \frac{\sigma_{\text{abs}}}{V} \right)_{\text{cs}} = \frac{f_V k}{L_1 - L_0} \text{Im} \left[ \frac{2/3 - L_0}{L_0 - \xi(\omega)} + \frac{L_1 - 2/3}{L_1 - \xi(\omega)} \right] \quad (59)$$

where  $f_V$  is the metal volume fraction,  $L_0 = 1/2 - 1/6\sqrt{1 + 8f_V}$  and  $L_1 = 1/2 + 1/6\sqrt{1 + 8f_V}$  are structural depolarization factors, and as in Sec. 3 we define  $\xi(\omega) = -1/\chi(\omega)$ . One can see that there are two quasistatic resonances that arise from coupling the plasmons at the interior and exterior interfaces. At visible and infrared frequencies, the material parameters of typical metals satisfy  $|\text{Re} \chi| \gg 1$ , such that a single resonance dominates the response and the volume fraction  $f_V$  of the metal must be small (as in Fig. 3). The cross-section is then given by

$$\left( \frac{\sigma_{\text{abs}}}{V} \right)_{\text{cs}} \approx \frac{2}{3} f_V k \frac{1}{\text{Im} \xi(\omega)} \approx \frac{2}{3} k \frac{|\text{Re} \chi(\omega)|}{\text{Im} \chi(\omega)} \quad (60)$$

where we see that the absorption cross-section is proportional to  $|\text{Re}\chi|/\text{Im}\chi$  instead of  $|\chi|^2/\text{Im}\chi$  due to the small optimal volume fraction, which decreases as  $\text{Re}\chi$  increases. There is an additional 33% reduction due to the second “material” resonance [45] at  $L_1$ . Additionally, tuning the resonant wavelengths requires very large core-shell thickness ratios, roughly 25:1 at  $\lambda = 1000\text{nm}$ , whereas at the same wavelength the optimal ellipsoids have aspect ratios of less than 10:1.

### E. Propagation length of a plasmon near the surface-plasmon frequency

It is sometimes stated [33, 176–179] that for any frequency at which  $|\text{Re}\epsilon_m| \gg \text{Im}\epsilon_m$ , for metallic permittivity  $\epsilon_m$ , the complex surface-parallel wavevector for a plasmon at a metal–dielectric interface is given by:

$$\beta = \frac{\omega}{c} \left[ \sqrt{\frac{\epsilon'_m \epsilon_d}{\epsilon'_m + \epsilon_d}} + i \frac{\epsilon''_m}{2(\epsilon'_m)^2} \left( \frac{\epsilon'_m \epsilon_d}{\epsilon'_m + \epsilon_d} \right)^{3/2} \right], \quad (61)$$

where  $\epsilon_d$  is the dielectric permittivity and we use single and double primes to denote real and imaginary parts, resp. However, this expression is only an approximation, and it is only accurate at low frequencies where the plasmon confinement to the surface is *not* highly subwavelength. In particular, it is not accurate near the resonant frequency at which  $\epsilon'_m \approx -\epsilon_d$ , even if  $|\epsilon'_m| \gg \epsilon''_m$ . The frequency-dependent propagation length of the plasmon is given by:

$$L_{\text{prop}} = \frac{1}{2\text{Im}\beta}, \quad (62)$$

which according to Eq. (61) would go to zero at  $\epsilon'_m = -\epsilon_d$ , even for a lossy medium (as is claimed in Ref. [177]). This is not correct.

The exact dispersion relation for a surface plasmon is [33]

$$\beta = \frac{\omega}{c} \sqrt{\frac{\epsilon_m \epsilon_d}{\epsilon_m + \epsilon_d}}. \quad (63)$$

The approximation in Eq. (61) arises from a Taylor expansion of the imaginary part of  $\epsilon_m$  in the denominator of Eq. (63). However, such a Taylor expansion is only valid if  $|\epsilon'_m + \epsilon_d| \gg \epsilon''_m$ . Thus we see that this approximation fails near the surface-plasmon frequency, which is the exact regime where significant subwavelength confinement is possible. If we define the surface plasmon frequency  $\omega_{\text{sp}}$  such that  $\epsilon'_m(\omega_{\text{sp}}) = -\epsilon_d$ , then the dispersion relation is approximately

$$\beta(\omega_{\text{sp}}) \approx \frac{\omega}{c} \frac{\epsilon_d}{\sqrt{\epsilon''_m}} \left( \frac{1+i}{\sqrt{2}} \right), \quad (64)$$

assuming  $|\epsilon'_m| \gg \epsilon''_m$  (as above). It is then straightforward to derive the propagation length at  $\omega = \omega_{\text{sp}}$ :

$$L_{\text{prop}} \approx \frac{\lambda}{2\sqrt{2}\pi} \frac{\sqrt{\epsilon''_m}}{\epsilon_d}, \quad (65)$$

which does not go to zero except in the case of zero loss. (In the case of zero loss the group velocity also goes to zero, with the same dependence on  $\epsilon''_m$ , such that the decay time is nonzero and finite. As the material loss goes to zero, the energy loss per unit time remains nonzero but finite, because the resonant field enhancement balances the vanishing material loss.) Thus near the surface-plasmon frequency the propagation length is proportional to  $\sqrt{\text{Im}\epsilon_m}$ , not  $(\text{Re}\epsilon)^2/\text{Im}\epsilon$  (which is the scaling of Eq. (61) and which only applies in the low-frequency Sommerfeld–Zenneck regime).

## F. Optimal LDOS of a thin film

We derive Eq. (43), the optimal LDOS of a thin film at frequencies  $\omega \ll \omega_{\text{sp}}$ . We measure the LDOS a distance  $d$  from the thin film, which we take to have a thickness  $t$ . The LDOS of a thin film, relative to the free-space LDOS  $\rho_0$ , is [69]

$$\frac{\rho}{\rho_0} \approx \frac{1}{k^3} \int k_p^2 e^{-2k_p z} (\text{Im} S_{21}) dk_p \quad (66)$$

where  $k_p$  is the magnitude of the surface-parallel wavevector, and we have assumed the primary contribution to the integral comes from high-wavevector (p-polarized) waves for which  $e^{2ik_z d} \approx e^{-2k_p d}$ . For small imaginary permittivity  $\text{Im} \varepsilon \ll |\text{Re} \varepsilon|$ , which is a requirement to access the large wavevectors,  $\text{Im} S_{21}$  is a sharply peaked function centered around a parallel wavevector  $k_{p0}$  that depends on the thin-film thickness  $t$  by [95]

$$k_{p0} \approx \frac{2}{|\varepsilon| d} \quad (67)$$

which we will use below to determine the optimal thickness. The integral can be approximated by

$$\int k_p^2 e^{-2k_p z} (\text{Im} S_{21}) dk_p \approx \frac{\pi}{2} k_{p0}^2 e^{-2k_{p0} z} [\text{Im} S_{21}]_{\text{max}} \Delta k_p \quad (68)$$

where we have assumed that  $\text{Im} S_{21}$  is a Lorentzian with full-width half-max of  $\Delta k_p$ . As derived in Ref. [95], the ‘‘reflectivity–bandwidth product’’ is

$$[\text{Im} S_{21}]_{\text{max}} \Delta k_p \approx 2k_{p0} \quad (69)$$

Thus the LDOS enhancement is proportional to  $k_{p0}^3 e^{-2k_{p0} z}$ , which reaches a maximum for a thickness such that  $k_{p0} = 3/2z$ , yielding an optimal LDOS of

$$\begin{aligned} \left[ \frac{\rho}{\rho_0} \right]_{\text{max}} &\approx \frac{27\pi}{8e^3} \frac{1}{(kd)^3} \\ &\approx \frac{1}{2(kd)^3} \end{aligned} \quad (70)$$

## G. Frobenius norm of the dyadic Green’s function

The dyadic Green’s function defined in Eq. (33) is given by [51]

$$\overline{\overline{G^{EP}}} = \frac{k^2 e^{ikr}}{4\pi r} \left[ \left( 1 + \frac{i}{a} - \frac{1}{a^2} \right) \delta_{ij} + \left( -1 - \frac{3i}{a} + \frac{3}{a^2} \right) \frac{x_i x_j}{r^2} \right] \quad (71)$$

where  $a = kr$ . The Frobenius norm is given by  $G_{ij}^{EP} \overline{G_{ij}^{EP}}$  (momentarily denoting complex conjugation with an overline) and can be computed by use of the identities (repeated indices indicating summation)

$$\begin{aligned} \delta_{ij} \delta_{ij} &= \delta_{ii} = 3 \\ \delta_{ij} \frac{x_i x_j}{r^2} &= \frac{x_i x_i}{r^2} = 1 \\ \frac{x_i x_j}{r^2} \frac{x_i x_j}{r^2} &= 1 \end{aligned}$$

which yield the expression in Eq. (33).

**Acknowledgments**

We thank Eli Yablonovitch, Adi Pick, Yehuda Avniel, Aristeidis Karalis, Hrvoje Buljan, Alejandro Rodriguez, and Emma Anquillare for helpful discussions. This work was supported by the Army Research Office through the Institute for Soldier Nanotechnologies under Contract No. W911NF-07-D0004, and by the AFOSR Multidisciplinary Research Program of the University Research Initiative (MURI) for Complex and Robust On-chip Nanophotonics under Grant No. FA9550-09-1-0704.