Optical Materials for Maximal Nanophotonic Response

Hyungki Shim,1, 2, 3 Zeyu Kuang,1, 2 and Owen D. Miller1, 2

1 Department of Applied Physics, Yale University, New Haven, Connecticut 06511, USA
2 Energy Sciences Institute, Yale University, New Haven, Connecticut 06511, USA
3 Department of Physics, Yale University, New Haven, Connecticut 06511, USA

This article reviews the material properties that enable maximum optical response. We highlight theoretical results that enable shape-independent quantification of material “figures of merit,” ranging from classical sum rules to more recent single-frequency scattering bounds. A key delineation at optical frequencies is between polaritonic materials that support highly subwavelength resonances and dielectric materials that can have vanishingly small loss rates. We discuss the key metrics that enable comparisons both within these material classes and between them. We discuss analogous metrics for 2D materials, and point to applications for which rigorous comparison can be made between bulk- and 2D-material approaches. The various results highlight the synergy between materials discovery and theoretical nanophotonic bounds, and point to opportunities in achieving new extremes in light–matter interactions.

I. INTRODUCTION

As material discovery proceeds at a rapid pace [1–21], and atomistic control promises the possibility of “designer” materials [22–25], there is a fundamental question to be answered: what material properties should the optical-materials community aim to synthesize? In the field of nanophotonics, for many applications the goal is to maximize the interaction of light with matter, manifested by absorption and/or scattering [26–31], quality factor [32–38], spontaneous-emission rate [39–42], and related response functions [43–46], for frequency bands ranging from the near ultraviolet to the far infrared. In this Review, we survey the key metrics that have been identified for maximum optical response. Across the broad landscape of polaritonic, dielectric, and 2D materials, we use experimental optical-constant data in tandem with these metrics to identify especially promising materials and material characteristics at optical frequencies.

II. POLARITONIC MATERIALS

Polaritonic materials at optical frequencies benefit from the strong coupling of light to free electrons, but at the cost of significant absorption. To what extent can absorption losses be avoided, for large scattering or high-Q resonances? In this section we highlight two re-
sults that answer this question: bounds on the largest single-frequency response possible (Sec. II A) and bounds on the highest quality factor possible (Sec. II B), yielding two metrics (|χ|^2/Im χ and ∂(Re χ)/∂ω)/Im χ, respectively, for bulk materials with susceptibilities χ by which all polaritonic materials can be compared. We also highlight the important role the real part of the permittivity plays to determine the feasibility of achieving high-confinement polaritonic resonances (Sec. II C). In Sec. II D we highlight general bounds for maximum response in a regime of high-radiative-efficiency plasmonics, and hybrid dielectric-metal structures that offer a combination of high efficiency and large response, approaching their respective bounds. Finally, we discuss the important role nonlocality plays at small length scales, and known bounds that incorporate the relevant nonlocal parameters (Sec. II E).

A. Maximal Single-frequency Response

It has long been recognized that reducing material “loss” is critical for many plasmonics applications, yet there are many possible loss rates to choose from: the imaginary part of the susceptibility, the imaginary part of the refractive index, the real part of a material’s conductivity, the inverse of its real resistivity, etc. One way to frame the question is in the context of scattering problems: given an external excitation, what is the largest possible response (absorption, scattering, etc.) from a given material? Ref. [51] develops a systematic answer to this question, independent of particulars of the geometrical patterning. The idea stems from consideration of the polarization currents P excited within the material. The rate at which energy is absorbed is proportional to \( \int \text{Im} \chi |E|^2 = \text{Im} \chi |\chi|^2 \int |P|^2 \), and is proportional to the square of the magnitude of the polarization currents. By contrast, the well-known optical theorem [65–68] dictates that the total extinction of the incident field, via absorption or scattering, must be proportional to the imaginary part of a scattering amplitude, which increases only linearly with the induced polarization currents. Extinction must be larger than absorption, and yet absorption increases more rapidly with the magnitude of the polarization currents. Constraining absorption to be smaller than extinction, which is equivalent to requiring that scattered power be non-negative, thereby imposes a limit on the largest polarization currents that can be excited in any particular material. This constraint can then be used to identify bounds on many optical-response functions. The exact form of the bound depends on the application, but within each bound there is a “material figure of merit” \( f(\omega) \) that dictates increased possible response based on the material properties. (A similar analysis [69] can be done for the induced surface currents in 2D materials with local surface conductivities \( \sigma(\omega) \).) For any polaritonic material, the material metric is given by

\[
f(\omega) = \begin{cases} \frac{|\chi(\omega)|^2}{\text{Im} \chi(\omega)} & \text{3D / bulk materials} \\ Z_0 \frac{|\sigma(\omega)|^2}{\text{Re} \sigma(\omega)} & \text{2D materials}, \end{cases}
\]

where the impedance of free space, \( Z_0 \), is included in the 2D-material FOM to make it dimensionless. This material metric determines the maximal single-frequency response for many applications, including absorption, scattering, and local density of states [51, 69, 70], cross density of states [59], near-field radiative heat transfer [71], high-radiative-efficiency plasmonics [72], free-electron radiation [73], Raman scattering [74], and more. Equation (1) is for scalar, nonmagnetic materials; more generally, with a tensor susceptibility \( \chi \) or conductivity \( \sigma \), the material metric becomes \( \| \chi^\dagger (\text{Im} \chi)^{-1} \chi \|_2 \) (Ref. [51]) and \( Z_0 \| \sigma^\dagger (\text{Im} \sigma)^{-1} \sigma \|_2 \) (Ref. [75]), respectively, where \( \| \cdot \|_2 \) denotes the matrix 2-norm [76].

Intuitively, the material metric of Eq. (1) makes sense: larger absolute susceptibilities imply the ability to sustain large currents, while the imaginary part of the susceptibility must dampen resonant response. Figure 1(a,b) compares different polaritonic materials against the material FOMs \( f(\omega) \) of Eq. (1) for experimentally characterized bulk and 2D materials. In the figure for bulk materials, across a spectrum ranging from the extreme UV to the mid-infrared, there is a clear trend for increasing \( f(\omega) \) with wavelength, which can be attributed to Drude-like response in such materials. For a Drude susceptibility \( \chi(\omega) = -\omega^2/\gamma \), the material FOM is given by \( \omega^2/\gamma \), and is therefore exactly proportional to wavelength. Variations from linear dependence thus represent non-Drude features in the material susceptibilities. The increasing material FOM with wavelength may be compensated by frequency-dependent constants in the response function; for example, the far-field scattering bounds [51] multiply the material FOM by a factor \( \omega/c \) that exactly compensates a linear increase with wavelength. As we discuss further in Sec. II C, the large values of \( |\chi|^2/\text{Im} \chi \) for noble metals at infrared frequencies may represent bounds that are not achievable in practice (due to infeasible synthesis requirements), in which case polar-dielectric materials, transparent conducting oxides, and doped semiconductors may all be viable alternatives. In the case of 2D materials, per Fig. 1(b), one can see that graphene with a large Fermi level appears ideal at photon energies below 1 eV, while 2D Ag, Al, and Au all perform very well at higher photon energies. In that figure we take the 2D-material limit from bulk properties of the metals; intriguingly, ab-initio calculations suggest that actual single-layer sheets of 2D materials may have significantly larger material FOM than their infinitely thin bulk counterparts [94].

The bulk-material figure of merit of Eq. (1) appears in other contexts as well. In Ref. [95], it is shown
FIG. 1: (a) Comparison of representative bulk, polaritonic (and/or lossy) materials via the material figure of merit $|\chi(\omega)|^2 / \text{Im} \chi(\omega)$. Conventional metals (Ag, Au, Al, Cu) [77] outperform alternative plasmonic materials (aluminum-doped ZnO (AZO) [60], Dysprosium-doped CdO (CdO:Dy) [78], SiC [79], TiN [80], ITO [81], doped InAs [82], n-type and p-type Si [83]) in the visible and infrared (albeit at the expense of large permittivity real parts). Drude material susceptibilities exhibit linear increases in material FOM as a function of wavelength, explaining the linear trend in the figure. (b) Comparison of various 2D materials by material FOM $Z_0 \left| \sigma(\omega) \right|^{-1} \left( \text{Re} \sigma(\omega) / \text{Im} \sigma(\omega) \right)$ (or $Z_0 \left| \sigma(\omega) \right|^{-1} (\text{Re} \sigma(\omega) / \text{Im} \sigma(\omega))$ in the scalar case). Here, we compare: graphene at different Fermi levels [84] (solid black lines) and magnetic-biasing [85] (dashed black line), AA-stacked bilayer graphene [86] (dark red), hBN [87] (green), MoS$_2$ [88] (purple), the anisotropic conductivity components of black phosphorus [63] (BP, pink and dark purple), and three 2D metals [89], Al (red), Ag (blue), and Au (gold). High-Fermi-level graphene and 2D silver offer the largest possible responses at infrared and visible wavelengths, respectively. The inset compares graphene at THz frequencies [90–92] to the topological insulator Bi$_2$Se$_3$ [93], which can have a surprisingly large FOM. For (b), reprinted with permission from Ref. [69], American Chemical Society.

that $|\chi|^2 / \text{Im} \chi$ is the key material metric determining a geometry-independent fundamental limit to propagation length in plasmonic waveguides. The essence of that derivation similarly approaches the problem as one of identifying the maximum possible induced polarization currents. (It has been suggested [96–100] that the propagation length of a surface plasmon on a planar interface follows an expression that is ultimately proportional to $(\text{Re} \varepsilon)^2 / \text{Im} \varepsilon$, which is very similar to the material metric. However, this is derived with a Taylor expansion that is invalid on resonance, as explained in Appendix E of Ref. [51]; the correct expression is in fact proportional to the $\sqrt{\text{Im} \chi}$, which decreases as loss decreases, because of the concomitant reduction in group velocity. But these arguments only apply to planar surfaces and patterned surfaces may more closely approach the bound of Ref. [95].) The inverse of the material metric has also been identified as the fundamental loss quantity to be minimized in metamaterial-based models [101].

Epsilon-near-zero materials [102, 103] exhibit intriguing phenomena such as distorted channels [104, 105], high-directivity emission [106, 107], and arbitrarily large phase velocities [108, 109], and offer the possibility for significant enhancement of nonlinear optical response [110–112]. But they do not offer any particular benefit for large linear response. At the frequency where the real part of the permittivity crosses zero, the material FOM simplifies to $[1 + (\text{Im} \chi)^2] / \text{Im} \chi$, which will tend to be significantly smaller than many of the values in Fig. 1(a), due to the modest magnitude of $\chi$.

An ideal polaritonic material has a purely real, negative permittivity with zero loss, in which case the material FOM diverges. A real, negative permittivity over a nonzero, finite frequency band is compatible with causality requirements (e.g., the Kramers–Kronig relations), and in theory ultra-low loss metals may be achievable in materials with artificially large lattice spacing or in designer organometallic compounds [113]. The bounds for any physical power-flow quantity cannot themselves diverge, and must be regularized by other effects in the presence of a lossless material. In the case of cross-sections per volume, the bounds can diverge as nonzero response is possible with arbitrarily small volumes. However, cross-sections themselves cannot diverge as radiative losses must become dominant as absorptive losses go to zero [114–116]. In the near field, there is not necessarily any radiative coupling; one form of regularization would be the breakdown of a local bulk susceptibility at
the large wavevectors that are accessible with very small loss. In either scenario, lower-loss materials towards such regularizations would represent improved response relative to the current state-of-the-art. From a microscopic perspective, two fundamental sources of loss are intrinsic quantum linewidth broadening [117] as well as inhomogeneous broadening, but typical metals exhibit higher losses than required by these sources and thus significant improvements may be possible [118].

There has been significant effort in mitigating losses in polaritonic materials [50, 60, 113, 126–128]. The synthesis techniques for a given material can dramatically affect the material FOM, as depicted in Fig. 2. High-quality films of Ag and Au can be deposited using focused-ion-beam lithography or other techniques [122]. However, it is shown that, up to a threshold, the thickness of the film will adversely increase the optical loss [124, 125]. The dielectric function of bulk silver is sensitive to environmental conditions and influenced by extrinsic effects, such as surface, impurity, and grain boundary scattering [123]. On the contrary, bulk gold is less sensitive to sample morphology and variation between different bulk measurements [77, 119, 121] are mostly likely caused by systematic errors [120]. There are other approaches to mitigating loss as well. One approach is to engineer the free-carrier concentration [129–137]. Another is to use gain media to compensate for loss [138–140]. Polar dielectrics supporting surface phonon-polaritons tend naturally to have lower losses and are good polaritonic materials at mid-infrared frequencies [49, 63, 141]. While all these approaches help reduce loss as measured by the imaginary part of susceptibility, they do not necessarily translate to an enhancement in the material FOM. For example, reducing the free-carrier concentration decreases the magnitude of the real part of susceptibility, in addition to decreasing its imaginary part, thus resulting in smaller material FOM. Thus a useful measure of loss is the inverse of the material FOM, \( \frac{1}{\text{FOM}} \), which is a frequency-dependent weighted average of \( \frac{1}{\gamma} \) times the \( \Gamma_n \) rates, \( \gamma = \sum \Gamma_n / \Gamma_n \). Then the quality factor is bounded below by a simple ratio of the resonant frequency \( \omega \) to \( \gamma_{\text{max}} \):

\[
Q \geq \frac{\omega}{2\gamma_{\text{max}}(\omega)},
\]

where the extra factor of 2 in the denominator arises from the definition of \( \gamma_{\text{max}} \) as a weighted average of half of the Drude loss rates. Technically, Eq. (3) does not require a quasistatic approximation, but it does require material losses to dominate relative to radiative losses.

These derivations can be adapted easily to 2D materials, by making the replacement \( \omega \chi(\omega) \rightarrow i\delta_S(x)\sigma(\omega) \), where \( \delta_S(x) \) is a delta function on the surface of the (not necessarily planar) 2D material. The analog of Eq. (2) for a 2D material is:

\[
Q = -\frac{\omega^2}{2} \frac{\partial}{\partial \omega} \left[ \frac{\text{Im } \sigma(\omega)}{\text{Re } \sigma(\omega)} \right].
\]

Beyond the quasistatic approximation, it is possible to narrow the linewidth beyond the expressions of Eqs. (2)–(4), through e.g. Fano resonances [144, 145] and near-field coupling [146, 147], but such effects must necessarily occur at larger size scales, without the highly subwavelength confinement available in the quasistatic limit.

Figure 3 plots the quality factor computed by Eq. (2) for a wide variety of materials. To compare loss rates at very different frequencies, in Fig. 3(a) the \( Q \) factor of representative materials is shown over a range of moderate values of the real part of their permittivities. Figure 3(b) shows the \( Q \) factors of many materials at the frequencies for which \( \text{Re } \varepsilon = -2 \), where a subwavelength sphere exhibits a surface-plasmon or surface-phonon resonance. The \( Q \) factor near these resonances can be expressed as \( \omega / \gamma \) where \( \omega \) refers to the resonance frequency. The \( Q \) factor tends to increase with wavelength because...
optical phonon lifetimes for polar dielectrics are typically orders of magnitude larger than those of plasmonic metals, which more than compensate for the (roughly an order-of-magnitude) reduction in resonance frequency. Doped semiconductors are intermediate between the two classes, with their loss rates several times smaller, but resonance frequencies slightly smaller, than their metal counterparts.

C. Real part of permittivity

Polaritronics in the mid-IR spectrum is rich with applications in sensing and selective thermal emission [148–150], given that a wide variety of molecules exhibit fundamental vibrational and rotational modes in the mid-IR and that blackbody emission peaks in this range for typical temperatures [49, 61, 151]. As discussed above, metals can have very large material FOM $f(\omega) = |\chi|^2 / \text{Im } \chi$ at such frequencies because the real parts of their permittivities tend to scale as $1/\omega^2$. However, achieving the corresponding scattering bounds may be unrealistic. Achieving polaritonic resonances in materials with large negative real permittivities may require difficult-to-fabricate feature sizes. As an example, for ellipsoids to have a quasistatic/plasmonic resonance at a particular frequency requires their depolarization factors $L$ to equal $\text{Re}(-1/\chi)$ (Ref. [51]), which requires increasingly large aspect ratios as $L = \text{Re}(-1/\chi) \rightarrow 0$. As an example, for a resonance at 5 $\mu$m wavelength in silver would require an aspect ratio greater than 50, whereas the highest aspect ratios fabricated to date are roughly 30 (Ref. [152]). It is possible to shift resonances to longer wavelengths without high aspect ratios by increasing their size, and correspondingly the radiative damping, but the confinement is reduced and the materials start behaving more like perfect conductors rather than plasmonic materials [82]. Large, negative real part of permittivities are also undesirable for certain transformation-optics based devices and applications. For instance, Ref. [153] has designed a non-magnetic, cylindrical cloak at optical frequencies that require the real part of permittivity of the constituent metal wires to have a similar magnitude to that of the surrounding dielectric.

Figure 4 compares the magnitude of $\text{Re } \varepsilon$ for various polaritonic materials in the visible and infrared spectrum. The shaded region corresponds to permittivity values whose real parts are between -2 and -16.9, which for nanorod-like ellipsoidal nanoparticles maps to aspect ratios from 1 to 5. This may be considered the region within which true quasistatic resonances can be achieved.

A subtlety that arises for high-loss materials is that the negativity of the real part of the permittivity no longer becomes the true delineation of whether a material supports quasistatic polaritonic resonances. In quasistatic electromagnetism, Maxwell’s equations simplify to Poisson’s equation for the quasistatic fields. Poisson’s equation can be transformed to a surface-integral equation [70, 154–158] for surface-charge configurations $\sigma$ at all material interfaces, a surface-integral equation that can be written in the form $\hat{K}\sigma - \Lambda\sigma = -s$, where $\hat{K}$ is the Neumann-Poincare operator [157–159] that is a Green’s-
FIG. 3: (a) Quality factor (using Eq. (2), valid under quasistatic approximation) for representative materials as a function of real part of permittivity. For Ag, optical constants were obtained from both Palik [77] and Johnson and Christy (J&C) [119]. (b) Comparison of quality factor (Eq. (2)) for different materials, ranging from polaritonic materials—metals, metal alloys, and doped semiconductors—to polar dielectrics. Reprinted with permission from Ref. [49], De Gruyter.

FIG. 4: Comparison of (negative) Re $\varepsilon$ of various polaritonic materials for a wide range of frequencies, ranging from visible to mid-IR. The pink shaded region shows the allowed range of Re $\varepsilon$ for realistic aspect ratio of 1 to 5 for a prolate spheroid. Given this range, conventional metals are ideal in the visible, whereas alternative plasmonic materials are suited for infrared spectrum. The extent to which the real part of permittivity of different materials responds to wavelength varies, with SiC demonstrating very small wavelength range over the pink shaded region and doped InAs the largest in the mid-IR.
function convolution operator, $\Lambda$ depends on the material susceptibility $\chi$ (assuming vacuum exterior, easily generalizable) via $\Lambda = 1/2 + 1/\chi$, and $s$ is a source term proportional to the incident field. The key aspect relevant to this discussion is that in an eigendecomposition of $\hat{K}$, the response will be maximal if the real part of $-1/\chi$ is in the range $[0, 1]$. This is the true condition for plasmonic-like response for a lossy material: $\Re(-1/\chi) \in [0, 1]$, generalizing the simple negative-permittivity condition for materials with nontrivial loss rates.

D. High-radiative-efficiency plasmonics

In the preceding sections, it was emphasized that one wants quasistatic resonances for maximum confinement in polaritonic response, and that coupling to radiative channels reduces such response. Yet for applications where high radiative efficiency is important, in areas such as far-field imaging [28, 160], photovoltaics [161, 162], and quantum nanophotonics [163], it is important to identify bounds that incorporate radiative-efficiency constraints. This problem is considered in Ref. [72], where it is shown that for both far-field scattering and near-field LDOS quantities, one can derive bounds with the additional constraint of high radiative efficiencies. In particular, the results of Ref. [72] show that imposing a minimum radiative efficiency $\eta_{\text{min}}$ (for hard-to-achieve radiative efficiencies above 50%) effectively reduces the maximum possible scattering response by a factor $\eta_{\text{min}}(1 - \eta_{\text{min}})/4$, in which case one can define the material FOM by

$$f(\omega) = \frac{\eta_{\text{min}}(1 - \eta_{\text{min}})}{4} \frac{\chi(\omega)^2}{\Im \chi(\omega)}.$$  \hspace{1cm} (5)

The new material FOM, given by Eq. (5), explicitly identifies the tradeoff in response that must be sacrificed to achieve high radiative efficiency.

The radiative-efficiency-constrained bounds enable comparison [72] of all-metal [164–168] and hybrid metal–dielectric [72, 169, 170] approaches to high-radiative-efficiency plasmonics. Shown in Fig. 5 is one set of findings from Ref. [72]: a hybrid silicon-on-silver resonator could have superior plasmonic properties to an all-silver resonator when radiative-efficiency constraints are included. Shown in Fig. 5(a–d) are the absorption and scattering cross-sections of silicon-only, silicon–silver, and silver-only optimized resonator designs, with the hybrid silicon-on-silver resonator showing the largest scattering cross-sections per volume. Moreover, as shown in Fig. 5(e), these resonators approach their scattering-efficiency bounds, including the metric of Eq. (5). Also shown in Fig. 5(e) are the bounds for silver-only and gold-only nanoresonators, which lie below the actual scattering performance of the designed silicon-on-silver structure (solid blue lines). This implies that there is no silver-only or gold-only approach that can ever do better than the designed silicon-on-silver structure, no matter how optimized the patterning is. This shows the power of such bounds: they convey the ability to survey a research field and rank-order certain approaches relative to each other. As shown in Ref. [72], the hybrid structures are also superior to metal-only structures for near-field spontaneous-emission enhancements at high radiative efficiency.

E. Nonlocal effects

Another important consideration is the effect of nonlocality in material susceptibilities for media synthesized at single-nanometer length scales, whereby the polarization currents induced at a point $x$ are related to the electromagnetic fields at another point $x$. Such effects both shift resonant frequencies [171–174] as well as dampen the maximal possible response [75, 175, 176]. Reference [75] considers the maximum response when the material susceptibility is described in a hydrodynamic framework, where the currents behave like fluids with a diffusion constant $D$ and convection constant $\beta$ (both real-valued), in which case the current is the solution of a convection–diffusion equation driven by the electric fields. In a quasistatic framework that is relevant at the length scales where such effects are important, one can show [75] that the cross-section bounds depend on a competition of two “rates”: the material FOM $f(\omega)$, and a second term relating the size of the scatterer to the “diffusion” length in the material. If one defines a radius $r$ of the smallest bounding sphere containing the scatterer, and a plasmonic diffusion length $\ell_D = \sqrt{\varepsilon D/\omega_p^2}$, for plasma frequency $\omega_p$, the maximum extinction cross-section per area of a 2D-material scatterer is given by [75]

$$\frac{\sigma_{\text{ext}}}{A} \leq \left[ \frac{Z_0 |\sigma_{\text{loc}}|^2}{\Re \sigma_{\text{loc}}} \right]^{-1} \left[ \left( \frac{\sigma_{\text{loc}}^2}{\ell_D^2} \right)^{-1} \right]^{-1},$$  \hspace{1cm} (6)

where $\sigma_{\text{loc}}$ is the local contribution to the conductivity. Equation (6) shows a dramatic reduction in response at size scales well below $\ell_D$, serving both to illuminate the effects of nonlocality as well as potentially serving as a means to extract the value of $\ell_D$ itself for any 2D material from experimental measurements. The analog of Eq. (6) for bulk materials is straightforward, replacing the first term with the material FOM $f(\omega)$ from Eq. (1) and the second term from a volume integral instead of a surface integral. An intriguing next step would be to consider the bounds that are possible in ab-initio material models.

III. HIGH-INDEX DIELECTRIC MATERIALS

Dielectric materials, with $\Re \varepsilon > 0$, can have very small material loss rates, and their dominant loss channel is typically radiation. The positive real part of the permittivity prevents such materials from supporting qua-
FIG. 5: Schematic of hybrid dielectric–metal resonator, as shown in the top left. Scattering and absorption cross sections of (a) Si cylinder in free-space and (b,c) Si and Ag cylinders, respectively, above a semi-infinite Ag substrate with gap thickness $g = 2$ nm. Geometrical parameters (insets) are chosen to align their resonant wavelengths at 700 nm. The three structures are all illuminated by normally incident plane waves. In (b,c), the absorption includes the dissipation in both the particle and the substrate. (d) The dielectric–metal structure shows the highest per-volume scattering cross-section, because it simultaneously achieves large scattering cross-section $\sigma_{\text{sc}}$, high radiative efficiency $\eta$, and small particle volume $V$. (e) In the visible regime, the scattering capabilities of metal–metal geometries (Ag–Ag and Au–Au bounds), free-space metallic (Ag bound), and free-space dielectric (Si free-space) scatterers all fall short when compared with the dielectric–metal (Si–Ag) scatterer, which also approaches its own upper bound. For the Si–Ag and Ag–Ag structures, the gap size is fixed at 5 nm; the cylinder (both Si and Ag) height $h$ ranges from 40 to 60 nm in order to tune the resonant wavelength. Reprinted with permission from Ref. [72], American Chemical Society.

 sisstatic, highly subwavelength resonances (though very small mode volumes are possible [177–180]), and it elevates the importance of geometrical patterning in their response. Independent of the patterning, however, there is an important geometry-independent sum rule governing their response and identifying high refractive index as a key metric for dielectric materials.

In a homogeneous medium with refractive index $n$, the density of plane-wave electromagnetic states is proportional to $n^3$ and $\omega^2$ at frequency $\omega$ (Ref. [181]). If we consider an electric dipole radiating at a any point in the medium, it will efficiently couple to half of these states (effectively the half with the same polarization). The power that it radiates will be directly proportional to the density of states, and is called the (electric) local density of states, LDOS [182]. The LDOS is a measure of the relative energy density of modes at a given point in space relative to the total density of modes [183]. Denoting the free-space electric LDOS by $\rho_0$, it is given by the expression [184]

$$\rho_0(\omega) = \frac{n^3\omega^2}{2\pi c^3}.$$  \hspace{1cm} (7)

Now consider a dipolar source in a structured medium of refractive index $n$. For a point source at some position $\mathbf{x}$, the LDOS $\rho(\mathbf{x}, \omega)$ is highly dependent on the position $\mathbf{x}$, the frequency $\omega$, and the material and structuring. But, if one considers a certain integral over all frequencies (and discard a near-field coupling term that typically, though not always, corresponds to emission into material-absorption pathways), Barnett and Loudon showed [56] that there is a simplifying sum rule from causality arguments. They considered a source point in vacuum, with a structured medium surrounding it, and showed that the integral over all frequencies of the relative difference between the structured-medium LDOS and the vacuum
The sum rule of Eq. (8) states that for any patterning, the total LDOS of a system cannot be modified, on average, over all frequencies. The LDOS can be increased in one frequency range, but thereby must be reduced in another. The only way to increase LDOS over all frequencies is by increasing the refractive index $n$ of the medium itself, which increases $\rho_0(\omega)$ by $n^3$ per Eq. (7). This highlights the key role that refractive index can play in maximizing light–matter interactions.

The role of refractive index to increase the available states has been recognized in a wide variety of systems. Large refractive index enables smaller mode volumes [177–180] and correspondingly large spontaneous-emission enhancements [185–187]. In a photonic-crystal cavity, a typical defect mode has minimum mode volume $\sim (\lambda/2n)^3$, in which case larger refractive index widens the bandgap and increases the spatial confinement of the mode [188–190]. In the quest for miniaturization of nanophotonic “building blocks,” plasmonic split-ring resonators tend to exhibit large absorption and can be difficult to fabricate and miniaturize at optical frequencies [191, 192]. High-index dielectric nanoresonators exhibit strong electric and magnetic resonances [193] and can help overcome such problems [194, 195]. Large refractive index contrast between the core and air helps to confine the light within the high index medium. As a result, the reduced radiation loss improves the overall $Q$ factors of the resonator [55]. High refractive index is also necessary for the miniaturization of such nanoresonators [196], since the magnetic resonance of a sphere occurs at a wavelength $\lambda \approx nd$, where $n$ is the refractive index and $d$ is the diameter of a sphere [194, 195, 197]. With nanoresonators as building blocks, a high-index material with positive permittivity and unit permeability can create metamaterials with effective permittivities and permeabilities across the four possible quadrants [198–200]. Nanoresonators with electric- and magnetic-dipole resonances can be used to tailor scattering profiles [54, 201–207]; as one example, effective negative index can be used for Huygens’ metasurfaces with no reflection loss and tailored forward scattering [208, 209]. Alternatively, with only a single dipolar resonance, one can create perfect electric reflectors with near-unity reflection due to the negligible loss of the dielectric material [210, 211]. Finally, magnetic reflectors feature maximum electric field at the interface, dramatically enhancing light–matter interaction on that plane [212, 213].

A classical application of the density-of-states enhancement associated with high index is for absorption enhancement in solar cells. Random surface texturing on thick films with plane-wave-like states enables full occupation of all of the internal states, an $n^3$ enhancement relative to external illumination from all angles. The intensity is enhanced in proportion to the product of the density of states with the wave speed, $c/n$, ultimately yielding the $4n^2$ Yablonovitch limit to absorption enhancement [214, 215]. Wavelength-scale patterning has enabled much thinner structures to approach, though generally not surpass, the $4n^2$ limit [41, 216–223].

Figures 6 and 7 show the refractive indices of many common high-index materials over transparency windows at visible and infrared frequencies. Many of the materials that are transparent in the visible and near-IR are polar dielectrics that support phonon polaritons, and thus are not transparent in the mid-infrared. Conversely, many (though not all) of the materials transparent in the mid-IR are not transparent in the visible/near-IR. One can see in both figures that the mid-IR-transparent materials tend to exhibit larger values of $n$ relative to visible/near-IR-transparent materials, consistent with various models showing that the refractive index for semiconductors tends to increase with smaller energy gaps [196, 247–250]. One tradeoff that tends to come with higher refractive index is increased chromatic dispersion, as measured by the derivative of the refractive index with respect to wavelength, i.e., $dn/d\lambda$. Figure 7 clearly shows a nearly linear relationship between refractive index and chromatic dispersion. Thus an important materials-synthesis question is whether higher-index materials with small chromatic dispersion can be synthesized. Such materials would immediately improve the performance of a wide variety of dielectric metasurfaces [251–261].

IV. COMPARING POLARITONIC AND DIELECTRIC MATERIALS

The metrics reviewed in Secs. (II,III) are not compatible with each other, preventing straightforward comparisons between polaritonic and dielectric media. The bounds in Sec. II are inversely proportional to the loss rate of the material, which is practically 0 for many dielectric media. Conversely, the refractive-index sum rule of Sec. III does not account for the surface waves that are so important in polaritonic media. In this section, we highlight two measures that enable direct comparison of the two systems: total electron number for all-frequency response (Sec. IV A), and recently developed power–bandwidth metrics for any bandwidth of interest (Sec. IV B).
FIG. 6: Comparison of common high-index materials over transparent frequency ranges in the visible and IR spectrum. Curves with blue colors represent materials transparent over the visible and near-IR spectrum \((\text{MgF}_2 [224], \text{SiO}_2 [225, 226], \text{Al}_2\text{O}_3 [227], \text{Si}_3\text{N}_4 [228], \text{ZrO}_2 [229], \text{LiNbO}_3 [230], \text{GaN} [231], \text{ZnS} [232, 233], \text{TiO}_2 [234], \text{ZnSe} [235])\), and those with red colors over the mid-IR \((\text{Si} [236], \text{ZnTe} [237], \text{GaAs} [238], \text{Ge} [239], \text{InP} [240–242], \text{InAs} [243], \text{PbS} [242, 244], \text{Te} [245], \text{PbSe} [244], \text{PbTe} [246])\). The chromatic dispersion for the two classes of materials are measured at wavelengths of 500 nm and 5 \(\mu\)m respectively, with darker shades corresponding to higher values of chromatic dispersion.

FIG. 7: Dispersion characteristics of high-index materials transparent over the visible and infrared spectrum, as measured by the chromatic dispersion \(\frac{dn}{d\lambda} (\mu\text{m}^{-1})\) over representative visible and infrared wavelength of 500 nm and 5 \(\mu\)m respectively. Materials with higher refractive indices tend to be more dispersive.

A. Electron number

The total number of electrons \(N_e\) in a system is a key property to describe the maximum response of any material. Causality requires any linear electromagnetic response function to be analytic in the upper half of the complex-frequency plane [262], which enables contour-integral-based “sum rules” to connect response averaged over all frequencies to certain constants of the scatterer. In quantum systems, this leads to the well-known
Thomas–Reiche–Kuhn sum rule (or "f-sum" rule) [263–267], which relates the sum of oscillator strengths for energy level transitions to the electron number. One can apply this technique to the extinction cross-section of any optical scatterer, $\sigma_{\text{ext}}(\omega)$, yielding the sum rule [57, 58]:

$$\int_0^\infty \sigma_{\text{ext}}(\omega) \, d\omega = \frac{\pi \omega_p^2}{2e} V = \frac{\pi e^2}{2\varepsilon_0 m_e c} N_e \approx 1.67 \times 10^{-5} N_e \, [\text{m}^2\text{s}^{-1}],$$

(9)

where $\omega_p$ is the effective plasma frequency of the material and $V$ is the volume it occupies. Equation (9) says that the extinction of light summed over all frequencies is determined by the number of electrons in the scatterer, independent of its geometry and the incident-field polarization. In Ref. [58], extinction cross-sections are computed for different canonical geometries and materials including aluminum and silicon, verifying that the sum rule is indeed independent of the nanostructure and material platform. To maximize frequency-integrated extinction, it is advantageous to use materials with large electron number. However, Eq. (9) does not provide any guarantees about which frequency ranges will contain the resonances. The sum rule requires susceptibilities that satisfy Kramers–Kronig relations, diminishing to zero at high frequencies. The decay-to-zero requirement, though physically reasonable, means that even "dielectric" media (semiconductors, insulators, etc.) have a plasma-like response at large enough frequencies. Such response contributes to integrated extinction, often in a large way due to the negative susceptibility. This obscures the behavior of, for example, a transparent dielectric at optical frequencies, by accounting for transitions that occur at UV and X-ray frequencies and which can be the dominant contributor to the response of Eq. (9).

B. Maximal response over nonzero, finite bandwidth

While the Thomas–Reiche–Kuhn sum rule provides information on integrated optical cross section, it does not provide spectral information over any finite bandwidth and often overestimates dielectric-material interactions as described above—the sum rule weighs the entire spectrum equally instead of isolating particular frequencies of interest. To remedy these shortcomings, Ref. [59] establishes power–bandwidth limits by combining causality principles, which underlie the sum rules discussed in Sec. III and Sec. IV A, with the energy-conservation principles underlying the single-frequency bounds of Sec. II. The resulting bounds enable comparisons between polaritonic and dielectric media, and yield the single-frequency and all-frequency bounds in their respective asymptotic limits. The key is to connect the frequency-averaged response over a bandwidth $\Delta \omega$ around a center frequency $\omega_0$ to a single complex-valued frequency $\omega = \omega_0 + i\Delta \omega$ (Ref. [59, 178, 268]). (This simple form emerges for a Lorentzian average; other averaging "windows" can be used [59, 178].) Then one can derive bounds that depend on material figures of merit for which the material parameters $\chi$ or $\sigma$ are evaluated at this complex frequency, i.e. $\chi(\omega)$ or $\sigma(\omega)$. Then, for nonmagnetic materials with bulk susceptibilities $\chi(\omega)$ or 2D conductivities $\sigma(\omega)$, the material FOM is given by:

$$f(\omega = \omega_0 + i\Delta \omega) = \begin{cases} \frac{|\omega \chi|^2 + |\omega \chi| \Delta \omega}{|\omega| \Im(\omega \varepsilon)} & \text{3D / bulk materials} \\ \frac{|\sigma(\omega)|^2}{\Re \sigma(\omega)} & \text{2D materials}. \end{cases}$$

(10)

(For the general case of anisotropic, magnetic, and even spatially inhomogeneous media, see the Supplemental Material of Ref. [59].) It is evident from Eq. (10) that the material FOM over any nonzero bandwidth yields a finite value even for lossless materials (due to the imaginary part of $\omega$, given by $\Im(\omega) = \Delta \omega$), and thus enables comparison among all possible optical materials. The FOM for 2D materials has an identical functional form (albeit with a complex frequency) as the single-frequency 2D-material FOM, Eq. (1), whereas the functional form of the bulk-material FOM is now slightly more complex in Eq. (10) than its single-frequency counterpart in Eq. (1). As in the single-frequency case, the material FOM in Eq. (10) favors large $|\chi(\omega)|$ and small $\Im \chi(\omega)$ (and similarly for 2D materials), now evaluated at the complex frequency $\omega = \omega_0 + i\Delta \omega$. One can identify intuitive forms of the material FOMs for small bandwidth ($\Delta \omega \ll \omega_0$), for which the material FOM simplifies to:

$$f(\omega) \approx \begin{cases} \frac{|\chi(\omega)|^2}{\Im \chi(\omega)} & \text{lossy (e.g. polaritonic)} \\ \frac{\omega_0}{\Delta \omega} \frac{|\chi(\omega)|^2}{\Im \chi(\omega)} + 1 & \text{lossless (dielectric)} \\ \frac{|\sigma(\omega)|^2}{\Re \sigma(\omega)} & \text{2D materials}, \end{cases}$$

(11)

where the 2D material FOM retains its original simple form. For high-index lossless materials:

$$f(\omega) \approx \frac{\omega_0}{\Delta \omega} \chi(\omega) = \frac{\omega_0}{\Delta \omega} (n^2 - 1) \quad \text{lossless, high-index.}$$

(12)

The material FOM for metals in the small-bandwidth limit is dictated by material loss $\Im \chi(\omega)$, whereas the FOM for lossless materials (dielectrics) is dictated by relative bandwidth $\Delta \omega/\omega_0$. Lossy materials simplify to the single-frequency metric $|\chi|^2/\Im \chi$ because the bandwidth of their resonant response is dictated by material loss, and their single-frequency response can be maintained over the whole bandwidth $\Delta \omega$. The material FOM for lossless high-index materials is consistent with the metrics of Sec. III, where we saw that all-frequency response is determined by the refractive index; the material FOM of Eq. (12) increases in proportion to $\omega_0$ over $\Delta \omega$, which can be interpreted as the possibility for resonant amplification of the typical $n^2$ enhancement in a resonator with...
FIG. 8: (a,d,g) Isocurves of material FOM for a Drude metal (with material loss rate $\gamma = 0.1 \omega_p$), a lossless dielectric (of susceptibility $\chi = 9$), and a Drude 2D material (with $\gamma = 0.01 \omega_p$). The arrows indicate increasing material FOM in each case. (b) Comparison of material FOM for various bulk metals and polaritonic materials (those supporting surface-phonon polaritons) / dielectrics, keeping the bandwidth-to-center-frequency ratio $\Delta \omega/\omega_0$ fixed to 0.1. Part (c) compares surface-phonon-polariton-supporting materials at mid-IR wavelengths. (e),(f) Comparison of material FOM for varying bandwidths relative to the center wavelengths of 1.55 and 10 $\mu$m. At very narrow bandwidths, dielectrics offer greater possible response than metals. (h,i) Comparison of material FOM for 2D materials for different choices of center wavelength and $\Delta \omega/\omega_0$. (2D Al, Ag, and Au properties derived from their bulk counterparts.) Reprinted with permission from Ref. [59], APS.

$Q$-factor given by $\omega_0/\Delta \omega$. No assumption of single-mode or quasistatic behavior is made in Eq. (11), which holds for any number of resonances as well as more complex phenomena such as Fano interactions [144, 145] and exceptional points [269, 270].

Figure 8 compares the material FOM for a large variety of materials at optical frequencies. On the left side of the figure is the material FOM for canonical material types: (a) a Drude metal, $\chi(\omega) = -\omega_p^2/\omega^2 + i\gamma \omega$, for plasma frequency $\omega_p$ and loss rate $\gamma$, (d) a lossless, constant-susceptibility $\chi(\omega) = 9$ material, and (g) a “Drude” 2D material, with conductivity $\sigma(\omega) = i\omega_p/\omega + i\gamma)$. One can see that these three material types show very different characteristic dependencies of their FOM on fre-
quency and bandwidth. The Drude-metal FOM is nearly independent of bandwidth for small-to-moderate bandwidths and increases with the center wavelength (of the frequency band of interest), $\lambda_0$. On the other hand, a dielectric with constant permittivity is independent of center wavelength, and highly dependent on the bandwidth. Finally, 2D Drude conductors are somewhere in between. Loss originates from both the material parameter $\gamma$ as well as the bandwidth, and the material FOM favors small bandwidth and large wavelength (for a large conductivity). These simplified permittivity/conductivity models describe the key features of the FOM for real materials, as the plots in Fig. 8(b,c,h) follow the same trends as those in Fig. 8(a,d,g): the material FOM for metals increases with wavelength, whereas dielectrics (Si and SiC) and polaritonic materials (SiO$_2$ and TiO$_2$) that support surface-phonon polaritons at mid-IR frequencies [47] are relatively constant with wavelength. Conversely, the plots in Fig. 8(e,f,i) show the effects of increasing bandwidth, with the FOM values nearly unchanged for metals but those of the dielectrics and polaritonic materials decreasing nearly linearly. The material FOM of 2D conductors increases with both wavelength and smaller bandwidths.

V. LOOKING FORWARD

The results highlighted in this paper demonstrate a synergy between experimental materials discovery and theoretical nanophotonic bounds. There are now a number of key metrics by which materials can be evaluated for optical performance. Looking forward, these results prompt new questions in a variety of directions. First, the metrics can drive new-material synthesis, whether for moderate-negative-permittivity polaritonic material with particularly small loss (as measured by $\text{Im}[\chi/\chi|^2]$) or for dielectric materials with particularly large refractive indices $n$ and small chromatic dispersion. From a theoretical perspective, a quantum-mechanical analysis may suggest constituent atoms or alternative approaches to achieving anomalously large material figures of merit, or they may provide insight through novel bounds on how large such metrics can be. There is also the question of which materials are optimal for quantum-photon applications, where high radiative efficiency (as discussed in Sec. II D) is important, and a number of considerations beyond maximum response may be desirable. This highlights another area for exploration—alternative nanophotonic metrics. Beyond maximum response, quantities such as nonreciprocal transmission, isolation, selectivity, and others may be desirable [271–274]; bounds on such response functions may introduce new material metrics for such scenarios. Active nanophotonic platforms offer another area for exploration, with metrics such as switching speed taking increased importance. These examples provide a glimpse at the fertile opportunity for better understanding of the extreme limits of light–matter interactions.

[1] J. Goldberger, R. He, Y. Zhang, S. Lee, H. Yan, H.-J. Choi, and P. Yang, Nature 422, 599 (2003).
[2] O. M. Yaghi, M. O’Keeffe, N. W. Ockwig, H. K. Chae, M. Eddaoudi, and J. Kim, Nature 423, 705 (2003).
[3] M. F. A. M. van Hest, M. S. Dabney, J. D. Perkins, D. S. Ginley, and M. P. Taylor, Applied Physics Letters 87, 032111 (2005).
[4] J. M. Macak, H. Tsuchiya, L. Taveira, S. Aldabargerova, and P. Schmuki, Angewandte Chemie - International Edition 44, 7463 (2005).
[5] K. S. Novoselov, D. Jiang, F. Schedin, T. J. Booth, V. V. Khotkevich, S. V. Morozov, and A. K. Geim, Proceedings of the National Academy of Sciences 102, 10451 (2005).
[6] A. K. Geim and K. S. Novoselov, Nature Materials 6, 183 (2007).
[7] A. Ayari, E. Cobas, O. Ogundadege, and M. S. Fuhrer, Journal of Applied Physics 101, 1 (2007).
[8] D. Paci´l, J. C. Meyer, C. Girit, and A. Zettl, Applied Physics Letters 92, 1 (2008).
[9] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature 452, 970 (2008).
[10] C. R. Dean, A. F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. L. Shepard, and J. Hone, Nature Nanotechnology 5, 722 (2010).
[11] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Physical Review Letters 105, 2 (2010).
[12] J. E. Moore, Nature 464, 194 (2010).
[13] M. Osada and T. Sasaki, Advanced Materials 24, 210 (2012).
[14] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, Nature Nanotechnology 7, 699 (2012).
[15] A. K. Geim and I. V. Grigorieva, Nature 499, 419 (2013).
[16] Z. K. Liu, B. Zhou, Y. Zhang, Z. J. Wang, H. M. Weng, D. Prabhakaran, S.-K. Mo, Z. X. Shen, Z. Fang, X. Dai, Z. Hussain, and Y. L. Chen, Science 343, 864 (2014).
[17] S. Borisenko, Q. Gibson, D. Evtushinsky, V. Zabolotnyy, B. B¨ochner, and R. J. Cava, Physical Review Letters 113, 027603 (2014).
[18] B. Q. Lv, H. M. Weng, B. B. Fu, X. P. Wang, H. Miao, J. Ma, P. Richard, X. C. Huang, L. X. Zhao, G. F. Chen, Z. Fang, X. Dai, T. Qian, and H. Ding, Physical Review X 5, 1 (2015).
