Anttu, Nicklas; Mäntynen, Henrik; Sorokina, Anastasiia; Turunen, Jari; Sadi, Toufik; Lipsanen, Harri

Applied electromagnetic optics simulations for nanophotonics

Published in:
Journal of Applied Physics

DOI:
10.1063/5.0041275

Published: 07/04/2021

Document Version
Publisher's PDF, also known as Version of record

Journal of Applied Physics

Applied electromagnetic optics simulations for nanophotonics

Cite as: J. Appl. Phys. **129**, 131102 (2021); https://doi.org/10.1063/5.0041275
Submitted: 21 December 2020 . Accepted: 14 March 2021 . Published Online: 06 April 2021

Nicklas Anttu, Henrik Mäntynen, Anastasiia Sorokina, Jari Turunen, Toufik Sadi, and Harri Lipsanen

Journal of Applied Physics TUTORIAL scitation.org/journal/jap

Applied electromagnetic optics simulations for nanophotonics

Cite as: J. Appl. Phys. **129**, 131102 (2021); doi: 10.1063/5.0041275
Submitted: 21 December 2020 · Accepted: 14 March 2021 ·
Published Online: 6 April 2021

View Online Export Citation CrossMark

Nicklas Anttu,[1,2,a)] Henrik Mäntynen,[1] Anastasiia Sorokina,[1] Jari Turunen,[3] Toufik Sadi,[4] and Harri Lipsanen[1]

AFFILIATIONS

[1]Department of Electronics and Nanoengineering, Aalto University, FI-00076 Aalto, Finland
[2]Physics, Faculty of Science and Engineering, Åbo Akademi University, FI-20500 Turku, Finland
[3]Institute of Photonics, University of Eastern Finland, FI-80101 Joensuu, Finland
[4]Engineered Nanosystems Group, School of Science, Aalto University, FI-00076 Aalto, Finland

[a)]**Author to whom correspondence should be addressed:** nicklas.anttu@aalto.fi

ABSTRACT

Nanophotonics—the science and technology of confining, guiding, and making photons interact with matter at the nanoscale—is an active research field. By varying the geometry and constituent materials, nanostructures allow precise control of the scattering of incident light and tailoring of emitted light. In this Tutorial, we outline the use of the Maxwell equations to model the optical response of nanostructures. This electromagnetic optics approach uses the refractive indices of the constituent materials and the geometry of the nanostructures as input. For most nanostructure geometries, analytical solutions to the Maxwell equations are not available. Therefore, we discuss varying computational methods for solving the equations numerically. These methods allow us to simulate the optical response of nanostructures, as needed for design optimization and analysis of characterization results.

I. INTRODUCTION

The progress in recent years on controlled fabrication of nanostructures has opened the door to precise control of both materials and geometry (see Fig. 1 for examples).[7–13] By proper design of such nanostructures, we can tailor the interaction with light, that is, photons, to a high degree and optimize the optical response for varying light-based applications. Thus, nanophotonics—the science and technology of confining, guiding, and making photons interact with matter at the nanoscale—is a thriving field in both academic research and industry.

However, it is often difficult to intuitively predict the optical properties of nanostructures. Optics modeling can thus be a valuable tool. For the optimization of geometry and materials for varying applications, modeling allows to scan the suitability of a far broader range of nanostructures than feasible with costly and time-consuming prototype fabrication. For complicated design problems, a broad range of advanced and powerful optimization techniques exist,[14] but to employ these methods, it is crucial to be able to efficiently model the optical response of the nanostructures. Furthermore, modeling can help in analyzing the origin of varying features observed in optical characterization as well as give access to optical properties that are not possible to directly measure, such as the distribution of light intensity in the interior of a nanostructure.

In this Tutorial, we give an overview to using electromagnetic optics, as described by the Maxwell equations, for modeling the interaction between light and nanostructures. To keep the scope concise, we focus on nanophotonics applications in the visible and near-infrared (NIR) range, spanning approximately from 400 to 1700 nm in wavelength (unless otherwise stated, we denote by wavelength the wavelength of light in vacuum). Note however that nanostructures are, in principle, of interest for a much broader range of photon energies, with applications spanning at least from x rays[15] to THz radiation.[16] Furthermore, we focus on the linear optical response of the nanostructures.

In Sec. II, we present a model for electromagnetic optics based on the Maxwell equations. The materials and geometry define the possible applications, whereas the intended application puts requirements on the materials and geometry (Fig. 2). Therefore, in Sec. III,

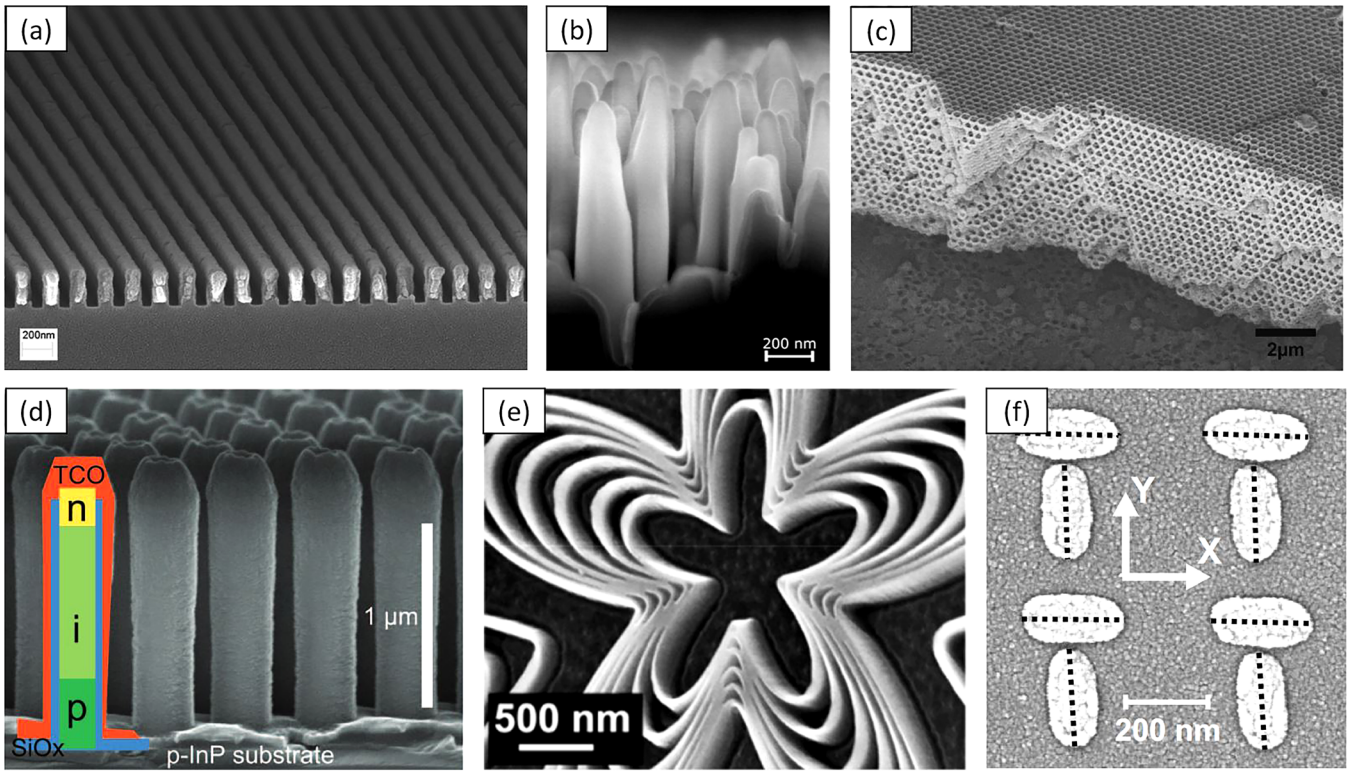

FIG. 1. Scanning electron microscope images of various nanostructures for photonics applications. (a) Al grating. Reproduced with permission from Kang *et al.*, Appl. Phys. Lett. **99**, 071103 (2011). Copyright 2011 AIP Publishing LLC. (b) Black-Si surface. Reproduced with permission from Pasanen *et al.*, Energy Procedia **124**, 307 (2008). Copyright 2017 Elsevier. (c) Inversed TiO_2 opal photonic crystal. Reproduced with permission from Guldin *et al.*, Nano Lett. **10**, 2303 (2010). Copyright 2010 American Chemical Society. (d) InP nanowire array solar cell with n, i, and p indicating the regions of the InP diode. Reproduced with permission from Wallentin *et al.*, Science **339**, 1057 (2013). Copyright 2013 the American Association for the Advancement of Science. (e) Nanopatterned Si substrate. Reproduced with permission from Khan *et al.*, ACS Appl. Nano Mater. **1**, 2476 (2018). Copyright 2018 American Chemical Society. (f) A group of Au nanodimers where the dashed lines indicate the slant of the nanostructures. Reproduced with permission from Husu *et al.*, Appl. Phys. Lett. **93**, 183115 (2008). Copyright 2008 AIP Publishing LLC.

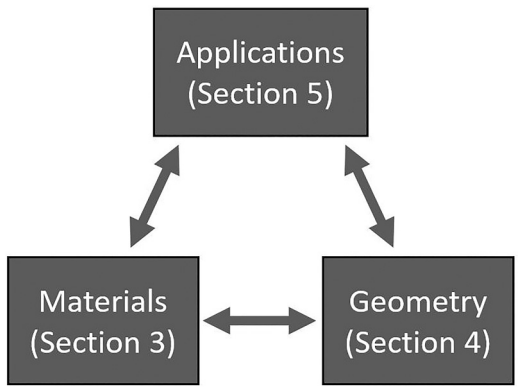

FIG. 2. The interplay in nanophotonics between intended application and available geometry and materials.

we introduce common types of materials and their optical properties; in Sec. IV, we give examples of popular nanostructure geometries in nanophotonics; and in Sec. V, we discuss applications of nanophotonics. In Sec. VI, we present briefly some of the popular analytical and numerical methods used in solving the Maxwell equations in nanostructures, and in Sec. VII, we reflect on the use of these varying simulation methods. In Section VIII, we comment on the impact of possible optical response beyond that within the model presented in Sec. II. Finally, we give concluding remarks in Sec. IX.

II. ELECTROMAGNETIC OPTICS AND THE MAXWELL EQUATIONS

Many optical phenomena, like mirror images and shadows, in our everyday life can be described with geometrical optics where rays of light are traced through the system. Such ray tracing can be successful also for designing lens-based optical instruments like telescopes and microscopes where the geometrical features of the

optical components are much larger than the wavelength of light.[17] However, the interaction of light with nanostructures is dominated by diffraction effects and then the wave nature of light must be taken into account in the analysis. Furthermore, the polarization state of light affects the diffraction. Therefore, we turn to an electromagnetic optics description of light where the Maxwell equations describe light as an electromagnetic field, which includes polarization through the vector nature of the field.

More specifically, to describe the scattering and emission of light, we use the macroscopic Maxwell equations (see Chap. 6.6 in Ref. 18),

$$\nabla \cdot \mathbf{B} = 0, \tag{1}$$

$$\nabla \cdot \mathbf{D} = \rho, \tag{2}$$

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}, \tag{3}$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0. \tag{4}$$

Here, $\mathbf{E}$ is the electric field, $\mathbf{D}$ is the electric displacement, $\mathbf{B}$ is the magnetic induction, $\mathbf{H}$ is the magnetic field, ρ is the free charge density, and $\mathbf{J}$ is the free current density.

We assume a time-harmonic form of the optical response: $\mathbf{E}(\mathbf{r}, t) = \mathrm{Re}[\mathbf{E}(\mathbf{r}, \lambda)\exp(-i\omega t)]$ and similarly for $\mathbf{D}$, $\mathbf{B}$, $\mathbf{H}$, ρ, and $\mathbf{J}$. Note that the physical field $\mathbf{E}(\mathbf{r}, t)$ is real-valued but $\mathbf{E}(\mathbf{r}, \lambda)$, which we use throughout for convenience, is complex-valued. Here, $\omega = 2\pi c/\lambda$ is the angular frequency with λ being the wavelength in vacuum and c being the speed of light in vacuum (and the photon energy is given by $E_{\mathrm{ph}} = \hbar\omega$, where $\hbar$ is the reduced Planck constant). For the material response, we assume that $\mathbf{D}(\mathbf{r}, \lambda) = \varepsilon(\mathbf{r}, \lambda)\mathbf{E}(\mathbf{r}, \lambda)$ and $\mathbf{B}(\mathbf{r}, \lambda) = \mu_0\mathbf{H}(\mathbf{r}, \lambda)$ with $\varepsilon(\mathbf{r}, \lambda) = n(\mathbf{r}, \lambda)^2\varepsilon_0$ being the permittivity, ε_0 being the permittivity of vacuum, n being the complex-valued refractive index, and μ_0 being the permeability of vacuum. Thus, we assume a linear, local, isotropic, non-magnetic, and time-harmonic optical response of the materials (see Sec. VIII for discussion of possible more complicated response types). In such a linear optics description, the material response is thus fully described through the wavelength-dependent refractive index.

In this model, $\mathbf{E}(\mathbf{r}, \lambda)$ and $\mathbf{H}(\mathbf{r}, \lambda)$ give full information of the electromagnetic field, from which, for example, the (complex-valued) Poynting vector $\mathbf{S}(\mathbf{r}, \lambda) = \frac{1}{2}\mathbf{E}(\mathbf{r}, \lambda) \times \mathbf{H}^*(\mathbf{r}, \lambda)$ is obtained. From $\mathbf{S}(\mathbf{r}, \lambda)$, the time-averaged Poynting vector can be calculated as $\mathrm{Re}(\mathbf{S}(\mathbf{r}, \lambda))$, and the time-averaged flow of intensity at position $\mathbf{r}$, in a direction given by the direction unit vector $\hat{\mathbf{n}}$, is given by $\hat{\mathbf{n}} \cdot \mathrm{Re}(\mathbf{S}(\mathbf{r}, \lambda))$. The integration of $\hat{\mathbf{n}} \cdot \mathrm{Re}(\mathbf{S}(\mathbf{r}, \lambda))$ over varying surfaces, with $\hat{\mathbf{n}}$ being the surface normal, can be particularly useful. For example, by integrating $\hat{\mathbf{n}} \cdot \mathrm{Re}(\mathbf{S}(\mathbf{r}, \lambda))$ over a closed surface, with $\hat{\mathbf{n}}$ pointing to the volume enclosed by the surface, we obtain the (net) absorbed power within the volume. In an emission problem, by placing the surface to encompass a non-absorbing volume surrounding the emitter, we obtain from the surface integration the total emitted power (when we chose $\hat{\mathbf{n}}$ to point outwards from the volume).

Importantly, dielectric, semiconductor, and metallic response, including absorption in the materials, can be taken into account through the wavelength dependence of the linear refractive index (see Sec. III for discussion of some of the widely used materials and their refractive indices). Such refractive index based modeling has reproduced excellently, for example, the measured diffraction and interference from absorbing semiconductor nanowire arrays, which show complicated dependence on geometry.[19]

A. Scattering of light

When we consider scattering of incident light, we set ρ and $\mathbf{J}$ to zero. Then, Eqs. (3) and (4) simplify to $\nabla \times \mathbf{E}(\mathbf{r}, \lambda) = i(2\pi c/\lambda)\mu_0\mathbf{H}(\mathbf{r}, \lambda)$ and $\nabla \times \mathbf{H}(\mathbf{r}, \lambda) = -i(2\pi c/\lambda)\varepsilon_0 n(\mathbf{r}, \lambda)^2\mathbf{E}(\mathbf{r}, \lambda)$ from which $\nabla \cdot \varepsilon_0 n(\mathbf{r}, \lambda)^2\mathbf{E}(\mathbf{r}, \lambda) = 0$ and $\nabla \cdot \mathbf{H}(\mathbf{r}, \lambda) = 0$, Eqs. (1) and (2), follow.[20,21] The incident light shows up as an external boundary condition to these equations.

B. Emission of light

To analyze the modification of emission by nanostructures, we consider a dipole emitter within or in the vicinity of the nanostructure. We include the dipole at position $\boldsymbol{r}_{\mathrm{s}}$ into the Maxwell equations through the source term [see Eq. (8.49) in Ref. 22],

$$\mathbf{J}(\mathbf{r}, \lambda) = -i\frac{2\pi c}{\lambda}\delta(\mathbf{r} - \mathbf{r}_{\mathrm{s}})\mathbf{p}, \tag{5}$$

where the dipole moment $\mathbf{p}$ defines the strength and orientation of the dipole emitter. Note that by the choice for $\mathbf{J}(\mathbf{r}, \lambda)$, we formally fix $\rho(\mathbf{r}, \lambda)$ since from Eqs. (2) and (3), $\nabla \cdot \mathbf{J}(\mathbf{r}, \lambda) = i\frac{2\pi c}{\lambda}\rho(\mathbf{r}, \lambda)$ follows.

This emission model includes the modified directionality and modified polarization of emitted light.[23] Furthermore, it allows the calculation of the Purcell factor C_{Purcell} (see Fig. 3 for example), which shows how much the nanostructure modifies, through diffraction effects, the optical (that is, radiative) recombination rate Γ_{Rad} of the dipole emitter [noting that the Purcell factor is directly proportional to the local density of optical states (LDOS) at the position of the emitter].[24] In other words, the Purcell factor tells how much the optical environment enhances (or suppresses) the optical recombination rate.[25] Thus, through the Purcell factor, the nanostructure can modify the emitter's internal quantum efficiency (IQE), which quantifies the fraction of recombination events resulting in the emission of a photon (with the rest of the recombination occurring through non-radiative recombination channels that do not give rise to the emission of photons but instead heat the sample).[26,27] Note that $\Gamma_{\mathrm{tot}} = \Gamma_{\mathrm{Rad}} + \Gamma_{\mathrm{NR}}$ is the total recombination rate, and the fraction of recombination events resulting in the emission of a photon is thus given by $\mathrm{IQE} = \Gamma_{\mathrm{Rad}}/(\Gamma_{\mathrm{Rad}} + \Gamma_{\mathrm{NR}})$.

To exemplify the effect of the Purcell factor on IQE, let us assume that the non-radiative recombination rate Γ_{NR} is not modified by the nanostructuring and that $\Gamma_{\mathrm{Rad},0}$ is the underlying optical recombination rate in the absence of nanostructuring. Thus, $\Gamma_{\mathrm{Rad}} = C_{\mathrm{Purcell}}\Gamma_{\mathrm{Rad},0}$ in the presence of the nanostructuring. Then, $\mathrm{IQE}_0 = \Gamma_{\mathrm{Rad},0}/(\Gamma_{\mathrm{Rad},0} + \Gamma_{\mathrm{NR}})$ is the original IQE, which is modified to $\mathrm{IQE} = \frac{\Gamma_{\mathrm{Rad}}}{\Gamma_{\mathrm{Rad}}+\Gamma_{\mathrm{NR}}} = \frac{C_{\mathrm{Purcell}}\Gamma_{\mathrm{Rad},0}}{C_{\mathrm{Purcell}}\Gamma_{\mathrm{Rad},0}+\Gamma_{\mathrm{NR}}} = \frac{C_{\mathrm{Purcell}}\mathrm{IQE}_0}{C_{\mathrm{Purcell}}\mathrm{IQE}_0+(1-\mathrm{IQE}_0)}$ by the nanostructuring.[26] Thus, with increasing C_{Purcell}, the probability

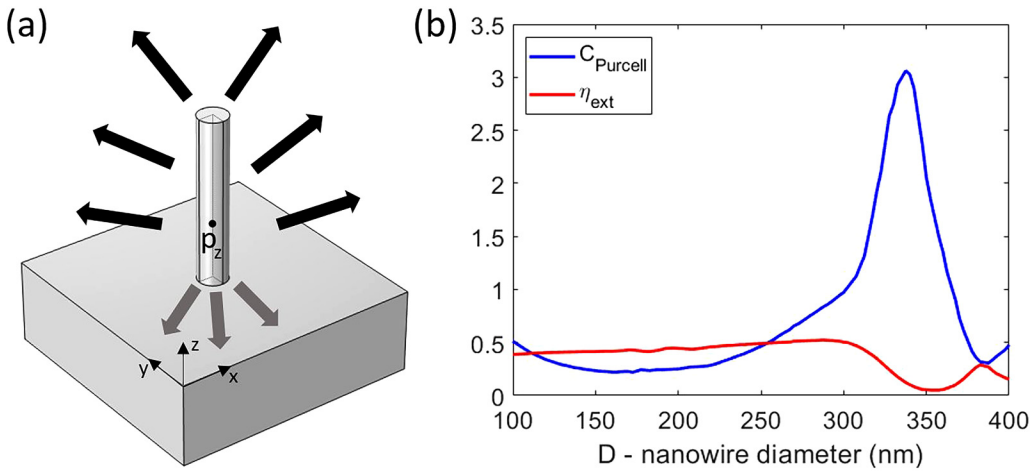

FIG. 3. (a) Schematics of a semiconductor nanowire on top of a semiconductor substrate, with a dipole at the center of the nanowire. The black arrows pointing upward indicate emission to the top side. The gray arrows pointing downward indicate emission into the substrate. (b) Simulated emission from a z oriented dipole ($\mathbf{p} = p_z\hat{\mathbf{z}}$) within a nanowire of length $L = 1000$ nm and varying diameter. The dipole is located at $z = 0.35L$ at the center of the x–y cross section of the nanowire and emits at $\lambda = 920$ nm. Here, we assumed a refractive index of $n = 3.5$ for the nanowire and the substrate, and $n = 1$ for the air top side. The modeling was performed with FEM; see Ref. 26 for technical details. We show (i) C_{Purcell}, which is the Purcell factor giving the modification in the optical recombination rate compared to the same dipole in a homogeneous semiconductor surrounding of $n = 3.5$, and (ii) η_{ext}, the extraction probability for an emitted photon. Here, η_{ext} is defined as the probability for the dipole to emit the photon to the top side [the black arrows in (a)].

for a radiative transition increases, and in the limit $C_{\text{Purcell}} \to \infty$, IQE $\to 1$ irrespective of the value of IQE_0. Conversely, with $C_{\text{Purcell}} \to 0$, IQE $\to 0$.

With this emission model, we can analyze the modified photoluminescence and blackbody radiation from extended emission regions within nanostructures, if incoherent dipoles are distributed throughout the emission region and their contributions summed to obtain the overall emission.[28] The spatially modified recombination rate can be used further as an input for light-emitting diode (LED) device modeling.[29]

Note that we study spontaneous emission with this model (modeling of stimulated emission would require in addition to the Maxwell equations, for example, the use of semiconductor Bloch equations[30]).

C. Plane waves and polarization of light

From the Maxwell equations, plane waves follow as a solution for a region where the refractive index n is constant. In other words, due to Eqs. (3) and (4), oscillating electric and magnetic fields sustain each other and allow for these types of propagating electromagnetic field solutions, even in the absence of sources. These plane waves are characterized by their propagation vector $\mathbf{k} = k\hat{\mathbf{k}}$ where $k = 2\pi n/\lambda = k_0 n$ and $\hat{\mathbf{k}}$ is the unit vector defining the propagation direction. The electric and magnetic field of the plane wave are of the form $\mathbf{E}(\mathbf{r}, \lambda) = \tilde{\mathbf{E}}(\lambda)\exp(i\mathbf{k}\cdot\mathbf{r})$ and $\mathbf{H}(\mathbf{r}, \lambda) = \tilde{\mathbf{H}}(\lambda)\exp(i\mathbf{k}\cdot\mathbf{r})$. Here, $\tilde{\mathbf{E}}(\lambda)$ defines the polarization of the light. From the Maxwell equations, $\mathbf{k}\cdot\tilde{\mathbf{E}} = 0$, $\mathbf{k}\cdot\tilde{\mathbf{H}} = 0$ and $\tilde{\mathbf{H}}\cdot\tilde{\mathbf{E}} = 0$ follow. Importantly, we can define two orthogonal polarization states with $\tilde{\mathbf{E}} = \tilde{\mathbf{E}}_1$ and $\tilde{\mathbf{E}} = \tilde{\mathbf{E}}_2$ such that $\tilde{\mathbf{E}}_1 \cdot \tilde{\mathbf{E}}_2 = 0$ (note that the magnetic field has corresponding polarization state given by $\tilde{\mathbf{H}} = \tilde{\mathbf{H}}_1$ and $\tilde{\mathbf{H}} = \tilde{\mathbf{H}}_2$). A general polarization state $\tilde{\mathbf{E}}$ of the plane wave can be expressed as a linear combination of any two such orthogonal polarization states.

In a system that contains a planar interface with surface normal $\hat{\mathbf{n}}$ (as, for example, the top surface of a substrate), it is common to define the so-called TE and TM (or alternatively s and p) polarization states of incident light. These polarization states are defined with reference to the incidence plane, which is the plane that contains both $\hat{\mathbf{n}}$ and $\hat{\mathbf{k}}$. TE (or s) polarization is defined as the polarization state where the electric field of the incident plane wave is perpendicular to this incidence plane (and the magnetic field of the plane wave lies then in this plane). Conversely, TM (or p) polarization is defined as the polarization state where the electric field of the incident plane wave lies in the incidence plane (and the magnetic field of the plane wave is then perpendicular to the incidence plane). In this way, TE and TM polarization denote which of the electric or magnetic field is transverse to the incidence plane, whereas s- and p-polarization denote whether the electric field is perpendicular (*senkrecht* in German, motivating the designation *s*) or parallel to this plane. Note that this is the most common definition for TE and TM, and s and p, polarization that we have encountered in the literature, but we recommend the readers to double-check the definitions whenever encountered in the literature. Also, when we discuss below, for example, x-polarized incident light, we mean light that has a non-zero electric field component only in the x direction (and similarly for y- and z-polarized light).

III. MATERIALS—WITH FOCUS ON THEIR WAVELENGTH-DEPENDENT REFRACTIVE INDEX

The refractive indices of the available materials set the foundation for the type of optical response that we can achieve in nanophotonics applications. To give general guidance in the options

available, we summarize the optical response of the three main types of optical materials typically used—insulators, semiconductors, and metals. For the exact choice of materials, the readers are referred to a detailed study of the refractive indices, for example, through summarizing collections[31] or online databases (e.g., www.refractiveindex.info) of experimentally measured values.

A. Insulators

The simplest materials for manipulating the scattering and diffraction of light could be considered insulators, e.g., oxides and polymers. These materials, which are often referred to as dielectrics, are governed in a large wavelength range (typically in an approximate range from 300 nm to 10 μm) by a real-valued refractive index. In this transparency wavelength range covering the visible and a large part of the IR region, the insulators show a real-valued, positive refractive index, typically in the range of 1.4–2.2, but some insulators like TiO_2 show a refractive index in the range of 2.5. There can be considerable dispersion in the refractive index, that is, variation of the refractive index with wavelength [see Fig. 4(a) for examples of refractive indices of insulators]. In the transparency region where Im

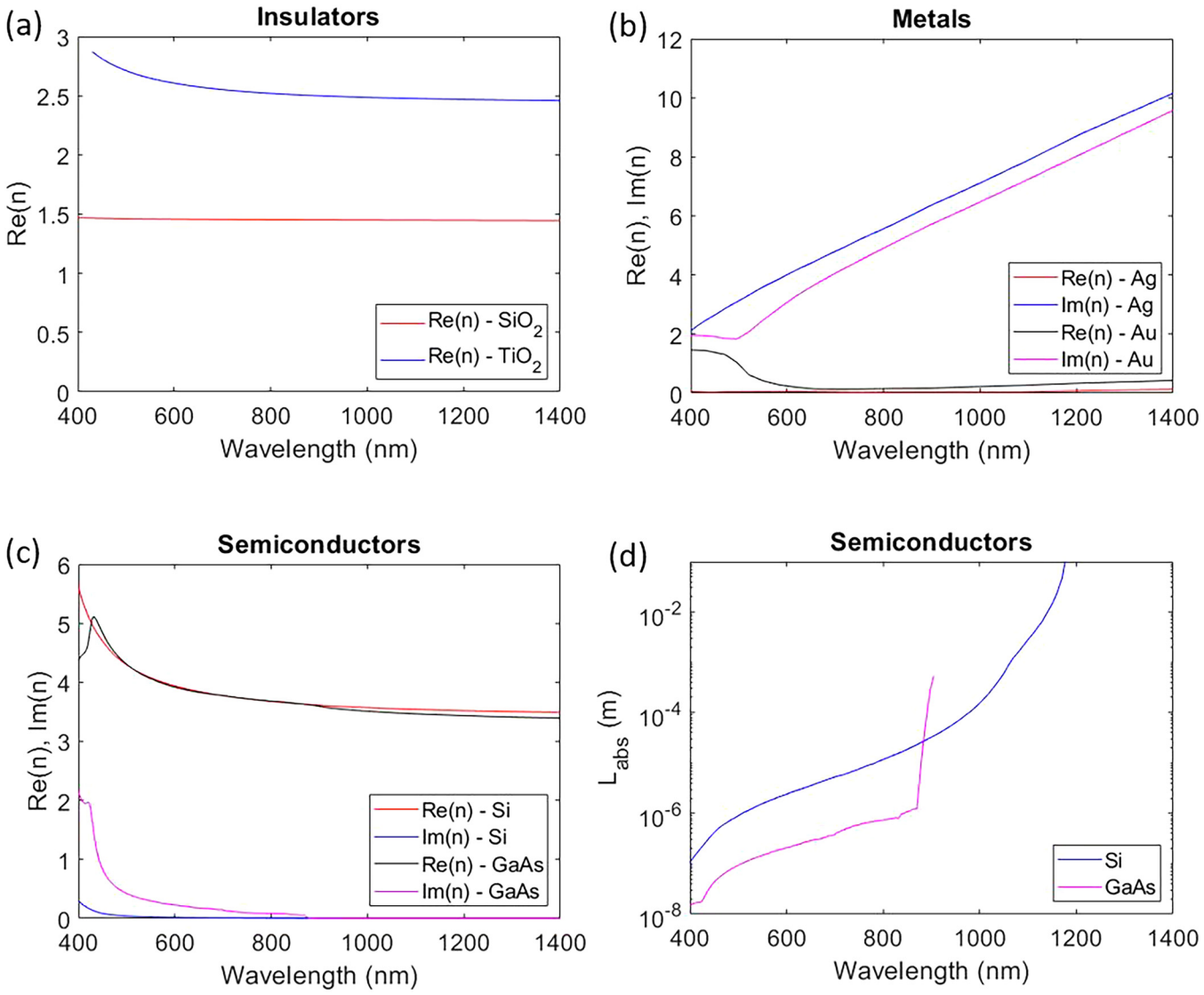

FIG. 4. (a) Refractive index of SiO_2 (fused silica) with values from Ref. 32 and TiO_2 with values from Ref. 33. In this wavelength range, Im(n) is negligible for these insulators. For TiO_2, which exhibits slight birefringence (in the form of a direction dependent refractive index defined by two components, n_o and n_e), we show values for n_o, the ordinary refractive index; the extraordinary refractive index n_e is ~0.2 higher. (b) Refractive index of Ag and Au, with values for both from Ref. 34. (c) Refractive index of Si with values from Ref. 35 and GaAs with values from Ref. 36. (d) Absorption length in GaAs and Si calculated from the Im(n) in (c).

$(n(\lambda)) = 0$, the real-valued refractive index can be considered to originate from the polarizability of the atoms constituting the material.[31]

For wavelengths below the transparency region, we observe Im $(n(\lambda)) > 0$. The absorption here occurs due to the excitation of valence electrons to the conduction band, that is, through interband transitions (for very short wavelengths in the x-ray range, we can also observe excitation of atomic core-states).[31] Similarly, for long enough wavelengths in the infrared region, the excitation of optical phonons can give rise to absorption, and the wavelength-ranges where such absorption shows up depend on the specifics of the material.[31] Importantly for this Tutorial, the wavelength range we consider is limited such that these two absorption mechanisms do not show up. Then, insulators can be considered non-absorbing materials with a real-valued refractive index $n(\lambda) = \mathrm{Re}(n(\lambda)) > 1$ [note that for vacuum, $n(\lambda) = \mathrm{Re}(n(\lambda)) = 1$, and for air the discrepancy from a value of 1 is typically below 0.1%].

B. Semiconductors

For optics, semiconductors differ from insulators in two major ways. First, the real part of the refractive index of semiconductors tends to be much higher than for insulators. For example, Si and III-V compound semiconductors show Re(n) in the range of 3–4.5 in the visible and NIR wavelength range [see Fig. 4(c)].[31] Therefore, semiconductors can, in principle, cause stronger diffraction of light than insulators.

Second, the interband transitions in semiconductors can occur in the visible or NIR range,[37] depending on the semiconductor. In other words, the bandgap energy E_{bg}, which defines the lowest photon energy that can cause interband transitions from the (mostly occupied) valence band to the (mostly empty) conduction band, is lower in semiconductors than in insulators. For convenience, we note that $E_{bg} \approx 1240/\lambda_{bg}$ with E_{bg} expressed in eV and λ_{bg} in nm, where λ_{bg} is the corresponding bandgap wavelength below which the interband transitions can occur (and $E_{ph} = \hbar 2\pi c/\lambda$).

One measure of absorption in semiconductors, and materials in general, is the absorption coefficient α (and its inverse—the absorption length L_{abs}): $\alpha(\lambda) = \frac{4\pi \mathrm{Im}(n(\lambda))}{\lambda}$. This coefficient is a measure of how far an electromagnetic plane wave can penetrate into a material [the absorptance A, that is, the fraction of original intensity that is absorbed, of a wave that has traveled a distance L inside the material is given by the Beer–Lambert law $A(\lambda, L) = 1 - \exp(-\alpha(\lambda)L) = 1 - \exp(-L/L_{abs}(\lambda))$]. For a low α at a given wavelength, light absorption is poor, leading, e.g., to a thin enough material layer appearing transparent at that wavelength. Semiconductors with a higher α absorb more readily photons, resulting in the photogeneration of excess electron–hole pairs, i.e., exciting electrons into the conduction band from the valence band, and hence creating holes in the valence band. In general, $\alpha(\lambda)$ is a practically useful quantity when choosing the material to be used to engineer solar cells and LEDs, for instance.

Note that since a photon with energy below the bandgap cannot excite electrons into the conduction band from the valence band, the absorption coefficient drops sharply at λ_{bg}, settling at practically zero for $\lambda > \lambda_{bg}$. At this point, it is worth distinguishing between direct bandgap (e.g., GaAs, InP, GaN) and indirect bandgap (e.g., Si, GaP) semiconductors.[37] In general, any photon absorption or emission event is subject to momentum conservation (or, equivalently, k-vector conservation). Note that the range of momentum values for photons in the visible and NIR range is negligible compared to the range of crystal momentum values for electrons in the semiconductor. Therefore, photons by themselves can induce transitions between electron states in the conduction and valence band at a fixed electron k-vector value.

In direct bandgap semiconductors, the conduction band minimum and valence band maximum occur at the same k-vector. Therefore, with $E_{ph} \approx E_{bg}$, a direct photon-mediated transfer (absorption or emission) of electrons between the valence and conduction band is possible. In contrast, in an indirect bandgap semiconductor, the valence band maximum and the conduction band minimum are offset in momentum space.[38] Then, in addition to the photon, also a phonon is needed in the transition for momentum conservation. Such requirement for an additional particle (the phonon) to participate in the optical transition typically drops the probability (and hence strength) of the transition by orders of magnitude. As such, a photon in an indirect bandgap semiconductor can typically penetrate much farther before its absorption, as compared to the direct bandgap case. This can be observed, for example, in Fig. 4(d), where the absorption coefficient for Si is considerably lower than that of GaAs (for $\lambda < 870$ nm, with 870 nm the λ_{bg} of GaAs). Note that at high enough photon energy, direct optical transitions become available also in indirect bandgap semiconductors: For example, in Si with $E_{bg} = 1.12$ eV, the direct transitions begin at $E_{ph} \approx 3.4$ eV.[38]

Importantly, for emission applications, the emission in bulk semiconductors occurs (usually) in the vicinity of λ_{bg} when excess electrons and holes, at the edge of the conduction and valence band, recombine in such a way that the excess energy is emitted through a photon. It is also worth noting that the emission, in terms of the IQE, is typically much higher in direct bandgap semiconductors compared to indirect bandgap semiconductors.

Throughout this Tutorial, we focus on room temperature applications, but we wish to note that the refractive indices of the materials are temperature dependent. For semiconductors, the temperature dependence of E_{bg} can cause noticeable effects where, as an example, λ_{bg} could shift readily by >5% if the temperature is varied by 200 K.[37] Such a shift of λ_{bg} would have a considerable effect on the transparency window of the semiconductor and the wavelength of emitted photons.

Similarly as for insulators, at long wavelengths, the excitation of optical phonons causes absorption in semiconductors; but again, in this Tutorial, we consider a wavelength range around the visible and NIR range where such absorption events are not observed.

C. Metals

In metals, a major part of the optical response can be understood to originate from light that interacts with the large density of electrons (of the order of $n_{carrier} = 10^{23}$ cm^{-3}) in the free electron cloud. In contrast to the underlying polarizability of the atomic lattice that dominates the Re(n) in insulators and semiconductors, and which occurs in phase with the incident light, giving rise to Re $(\varepsilon/\varepsilon_0) > 1$ and Re(n) > 1, the electron cloud in metals interacts out of phase with the incident light. Therefore, a perfect metal, which

shows a response only from excitation of the electron cloud without dissipation, will have $\mathrm{Im}(n) > 0$ and $\mathrm{Re}(n) = 0$, corresponding to $\mathrm{Re}(\varepsilon) < 0$ and $\mathrm{Im}(\varepsilon) = 0$.[34]

A plane wave incident at a planar interface between a dielectric and the metal will decay into the metal, similarly as inside a semiconductor as discussed above, with a constant given by $\alpha(\lambda) = \frac{4\pi \mathrm{Im}(n(\lambda))}{\lambda}$ [with accompanying decay length, which for metals is referred to as the skin depth: $L_{\mathrm{skin-depth}}(\lambda) = 1/\alpha(\lambda)$; note the similarity to the absorption length L_{abs} in semiconductors]. However, in contrast to the semiconductors, such exponential decay of the plane wave inside the perfect metal would not be associated with absorption, since $\mathrm{Im}(\varepsilon) = 0$ for the perfect metal and absorption requires $\mathrm{Im}(\varepsilon) > 0$.[39] Instead, perfect 100% reflection occurs. For real metals, $\mathrm{Re}(n)$ is not exactly zero since the excited electron cloud can dissipate energy to the lattice. Then, a small fraction of the incident light is absorbed because $\mathrm{Im}(\varepsilon/\varepsilon_0) = 2\mathrm{Re}(n)\mathrm{Im}(n) > 0$ [see Fig. 4(b) for examples of $n(\lambda)$ for metals].

In contrast to interband transitions, so important for semiconductors where they occur for $\lambda < \lambda_{\mathrm{bg}}$, the excitation of the electron cloud in a perfect metal shows instead a lower wavelength limit λ' below which the excitation of the electron cloud does not occur, with λ' proportional to $(n_{\mathrm{carrier}})^{-1/2}$; this corresponds to the plasma frequency, which is proportional to $(n_{\mathrm{carrier}})^{1/2}$.[34] In principle, when semiconductors are doped to give a large electron or hole concentration, also they can show such excitation of the free carrier cloud. However, typical free carrier concentrations in doped semiconductors are in the 10^{15}–10^{19} cm^{-3} range, giving a λ' which is located above the visible and NIR wavelength range of focus in this Tutorial.

A metal, such as Ag, whose optical response is dominated by the excitation of the electron cloud through the whole visible range shows a mirror-like, almost 100% reflecting, surface. In addition to absorption through the excitation of the electron cloud, interband transitions are also possible in metals. This additional absorption reduces the reflection of a metal surface at selected wavelengths. Such interband transitions give, for example, the yellowish color of Au.

D. Summary of the optical properties of insulators, semiconductors, and metals

For the visible and NIR wavelength range considered in this Tutorial (see Fig. 4 for examples), (1) most insulators show refractive index in the range of 1.4–2.0, with the exception of some high refractive index oxides such as TiO_2 with a refractive index of around 2.5. (2) Semiconductors can show a much higher $\mathrm{Re}(n)$ and hence a potentially stronger diffraction of light as compared to insulators. (3) Semiconductors show $\mathrm{Im}(n) > 0$ due to interband absorption that creates excess charge carriers (electrons in the conduction band and holes in the valence band). The absorption occurs for $\lambda < \lambda_{\mathrm{bg}}$, whereas for $\lambda > \lambda_{\mathrm{bg}}$, semiconductors behave as high refractive index insulators. The absorption gives rise to a wavelength-dependent absorption length for light entering a planar semiconductor sample, which can range from a few nanometers to hundreds of micrometers or more (depending on the choice of the semiconductor material). (4) Metals show a response that corresponds to a small $\mathrm{Re}(n)$ and a high $\mathrm{Im}(n)$. In this case, light that enters a planar metal sample shows exponential decay, but in contrast to semiconductors, this exponential decay is not associated predominantly with absorption. Instead, strong reflection occurs.

IV. GEOMETRIES FOR NANOPHOTONICS

In this Tutorial, together with the refractive index discussed in Sec. III, the geometry of the system at hand defines its optical response for given incident light or for emission from a given position in the system. Therefore, we give below a brief overview of some of the geometries popularly used in nanophotonics (see Fig. 5 for example schematics). In connection to describing the geometry, we give examples of some of the possible optical responses arising due to the geometry. We give more detailed examples of applications in Sec. V. For the geometries, we make a distinction between (a finite number of) nanoparticles and large-area structures (that could consist, in principle, of an arrangement of nanoparticles). Notably, both types of structures can give rise to geometry-tunable colors but due to different type of nanophotonic effects.[40]

A. Nanoparticles

Nanoparticles refer, in general, to arbitrarily shaped structures that are small compared to the considered wavelength of light. There are arguably two distinctions that can be made considering nanoparticles: whether the nanoparticles are metallic (plasmonic) or dielectric and whether a single particle [see Fig. 5(a)] or an arrangement of particles [e.g., an oligomer with a few particles; see Figs. 5(b) and 5(c)] are considered. These considerations have important implications on the physics, applications, and the choice of simulation methods. Note that the following notions of dielectric nanoparticles mostly apply to semiconductor nanoparticles as well.

1. Single nanoparticles

In metallic nanoparticles, collective oscillations of the free conduction electrons can be considered plasmons, i.e., quasiparticles and quanta of these oscillations at well-defined frequencies that depend on the geometry and the choice of metal.[42] Due to the subwavelength scale of a single metallic nanoparticle, the nanoparticle supports localized surface plasmons (LSP) in which the plasma oscillations extend across the entire particle. It is possible to excite a LSP with an incident light field when its frequency matches that of the oscillation. In a simplified picture, the coupling takes place such that the electric field of the incident light exerts a force that collectively displaces the conduction electrons while the exposed positively charged lattice of ionic cores provides a counter force that pulls the electrons back, thus creating a system akin to driving a damped mass-on-a-spring oscillator. Even such a simple model has found use in describing LSPs. In general, the LSP oscillations are inherently lossy as their energy dissipates to scattered electromagnetic fields and to heat [since $\mathrm{Im}(\varepsilon) > 0$ for any real, that is, non-perfect, metal].[42] In the case where the diameter of the metal nanoparticle is much smaller than the wavelength of light, the wavelength position of the LSP resonance is independent of the diameter of the nanoparticle and given by an electrostatic solution [see Fig. 6(a)].[42]

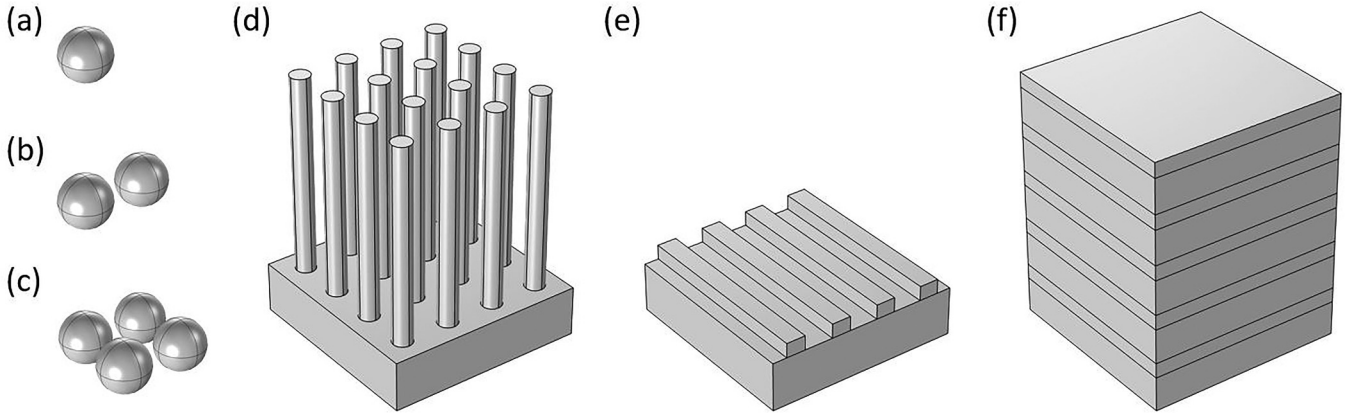

FIG. 5. Schematics of (a) a nanoparticle, (b) a nanoparticle dimer, (c) a nanoparticle tetramer, (d) a nanowire array, (e) a 1D grating, and (f) a planar optical stack.

LSP resonances in metallic nanoparticles exhibit two effects that are particularly relevant for applications. First, LSPs have a maximum optical extinction at the resonance frequency with an optical extinction cross section[41] that can be much larger than the geometrical cross section of the nanoparticle. Second, LSP resonances tend to focus the electromagnetic field close to the nanoparticle surface creating an electrical near-field with greatly enhanced field strength. Stronger and sharper resonances, as quantified by the so-called plasmonic quality factor, enhance both the extinction cross section and near-field strength. The LSP resonances depend on the properties of the nanoparticle itself, including size, shape, and material(s), as well as the dielectric environment and coupling to excitations of other nearby structures. It should also be noted that, depending on the nanoparticle size, some additional factors may need to be taken into account. For example, with large enough nanoparticle diameter, as compared to the wavelength of light, the LSP resonances redshift and may become multipolar and, with small enough nanoparticles, quantum effects may become prominent.[42]

Dielectric nanoparticles, on the other hand, can exhibit electric- as well as magnetic-type response of comparable strength

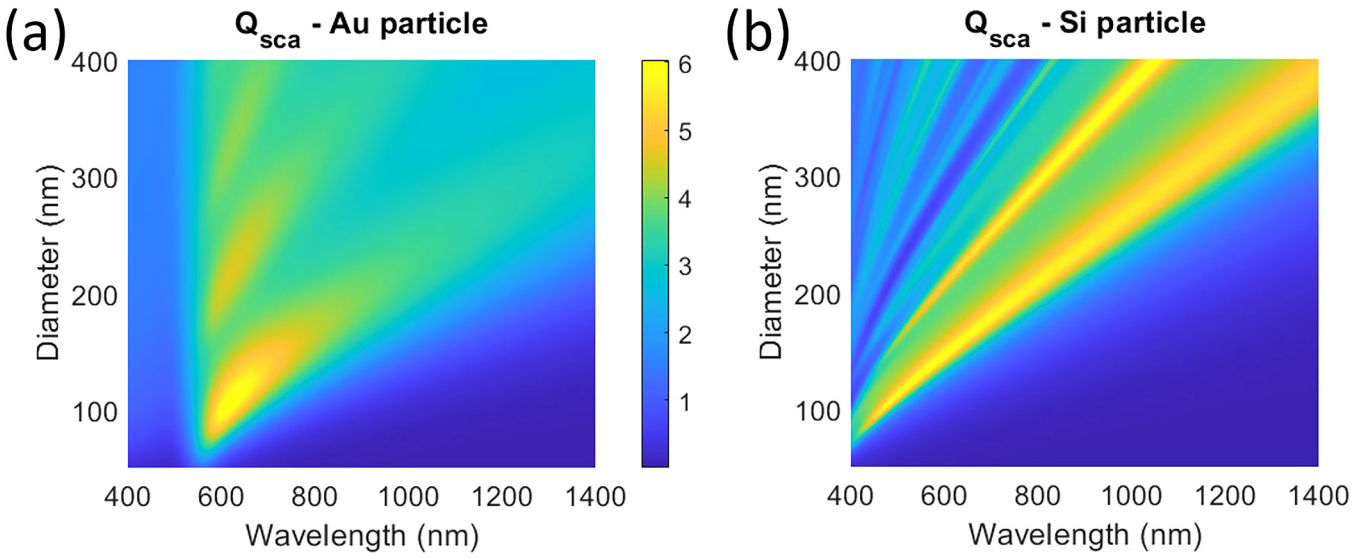

FIG. 6. Scattering efficiency Q_{sca} for varying diameter D of (a) an Au sphere and (b) a Si sphere in a surrounding of $n_{sur} = 1.5$. The calculations were performed with the Mie theory[41] for an incident plane wave using the refractive index of Au from Ref. 34 and Si from Ref. 35 for the refractive index n_{sph} of the sphere. Note that $Q_{sca} = P_{sca}/(I_{inc}A_{cs}) = P_{sca}/(I_{inc}\pi D^2/4)$ with P_{sca} being the power scattered by the particle, I_{inc} being the incident intensity of the plane wave, and A_{cs} being the cross-sectional area of the sphere. In (a), the LSP resonance settles toward a constant wavelength $\lambda_{LSP,small-D}$ with decreasing D. This wavelength is obtained by solving from the condition $\mathrm{Re}(\varepsilon_{sph}(\lambda_{LSP,small-D})) = -2\varepsilon_{sur}(\lambda_{LSP,small-D})$, which occurs at $\lambda_{LSP,small-D} = 529$ nm for the Au sphere in the surrounding of $n = 1.5$.[42]

and a much smaller energy dissipation to heat than in metallic nanoparticles [with no dissipation if we use $\mathrm{Im}(n) = 0$ for the nanoparticle material in the modeling]. The magnetic-type response can be understood to arise from the coupling of the incident electric field to circular displacement currents inside the particle. The resonance conditions for the electric and magnetic response can also have a different dependence on the nanoparticle geometry and, therefore, the resonances can be tuned separately.[43] In contrast to the LSP resonance in the above metallic nanoparticles, in dielectric nanoparticles, the resonance wavelength blueshifts continuously with decreasing diameter [see Fig. 6(b) for an example of a Si sphere].[41] If the dielectric or semiconductor is either non-absorbing or weakly absorbing in the wavelength range considered, essentially all of the energy in the incident electromagnetic field coupling to the nanoparticle will be re-radiated (scattered). Dielectric nanoparticles can also support resonant modes, which concentrate the fields inside the particle, which can be useful for absorption or nonlinear optics applications, such as third-harmonic generation.[44]

2. Nanoparticle oligomers

Nanoparticle oligomers ("oligo-" meaning "a few" in Greek) refer to arrangements of a few separate nanoparticles. The simplest oligomer is a dimer, i.e., two nanoparticles close to each other. Arrangements with more than two nanoparticles are usually regular rather than random. A typical example would be nanoparticles arranged at the vertices of a regular polygon or distributed along the circumference of a circle, possibly with an additional nanoparticle at the center. The nanoparticles also need not be similar and can exhibit different shapes and sizes or even different materials.

The collective excitations in oligomers can be thought to arise from the coupling of the excitations in each individual nanoparticle in analogy to the formation of molecular electronic orbitals from the valence electron wave functions of the individual atoms. Still following this analogy, the individual nanoparticle eigenmodes form symmetric and antisymmetric combinations, which hybridize to collective eigenmodes with increased or reduced eigenenergy for antibonding or bonding configuration, respectively.[45] The coupling strength is affected by the nanoparticle separation such that the coupling is stronger for smaller gaps between the particles and becomes negligible with a large enough separation. Therefore, the particle separation becomes an important extra design factor in addition to the properties of the particles themselves. In plasmonic nanoparticle oligomers, the LSP resonances couple to form collective resonant modes, and with close spacing (typically less than a few tens of nanometers), they tend to exhibit strongly enhanced electric fields in the gaps between the particles (see Fig. 7),[46] with the possibility of coupling also with resonances of the substrate.[47] The collective coupled resonances in dielectric nanoparticle oligomers, on the other hand, can exhibit strong magnetic hotspots in the gaps between particles.[43] However, the resonances can alternatively concentrate the fields inside the dielectric nanoparticles, in which case the coupling does not show such a strong dependence on the particle separation.[48]

The interaction of different resonances in nanoparticle oligomers can result in the so-called Fano resonances. Fano resonances are actually a more general concept, but in the context of optics, they arise due to the interference between an excited narrower

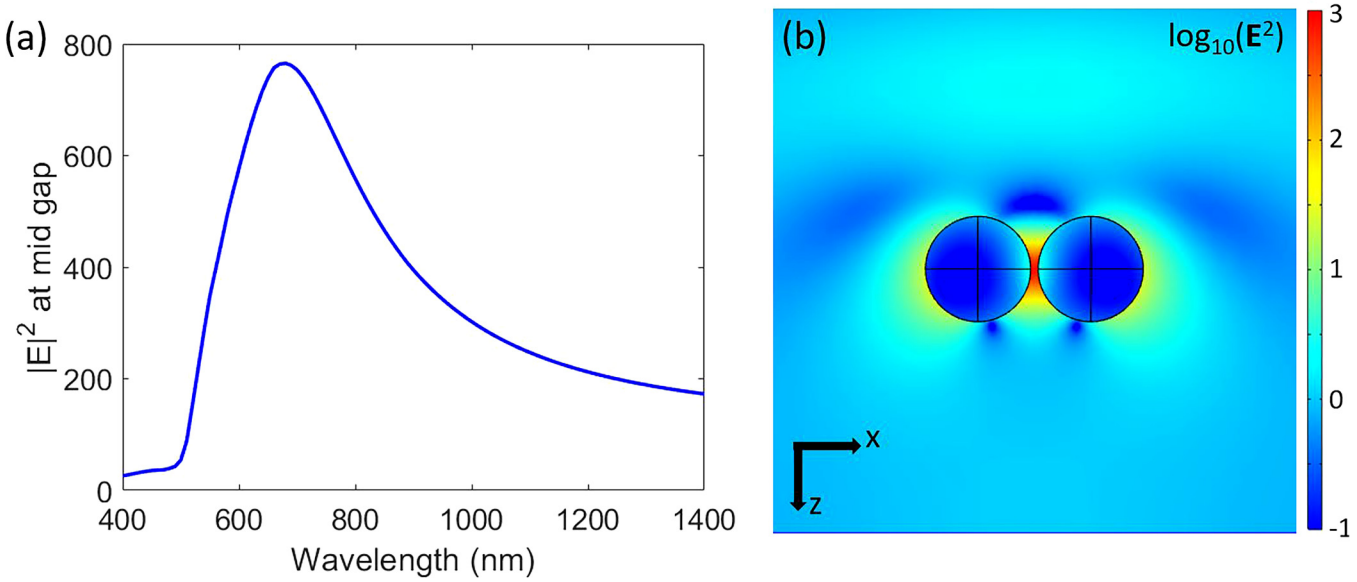

FIG. 7. (a) Electric field at the middle of the gap between two Au nanoparticles of 140 nm in diameter, separated by a gap of 10 nm, in a surrounding medium of $n = 1$. The axis of the dimer is in the x direction, and an x-polarized plane wave with $|E_{inc}| = 1$ V/m is incident along the z direction. (b) The electric field in the x–z plane through the center axis of the dimer in (a) at $\lambda = 680$ nm [that is, the peak in (a)]. The simulations were performed with FEM using the refractive index of Au from Ref. 34.

resonance and a broader resonant or nonresonant excitation in a nanostructure.[49] For example, a resonant magnetic dipole in a central, smaller Si nanoparticle interfering with the nonresonant background of the other Si nanoparticles in an oligomer can show a Fano resonance.[48] Similarly, the interference between super and subradiant modes in a plasmonic oligomer with in- and out-of-phase oscillations of the LSP resonance in the central nanoparticle can lead to Fano resonances.[46] Fano resonances are seen as features with a characteristic shape in the far-field spectra of absorption, transmission, or scattering.[49] It should be noted that there are also other resonant phenomena that can lead to similar spectral responses. The proper distinction between several such phenomena and Fano resonances has been recently addressed by Limonov *et al.*[49] It should also be noted that Fano resonances can occur in single nanoparticles as well.[43]

B. Large-area structures

Other common geometries employed in nanophotonics can often have a large, or theoretically infinite, extent even if they exhibit local nanostructuring with features smaller or comparable to the relevant wavelength of light. We restrict the following discussion to concern planar structures, which still covers a large scope of important applications.

1. Planar optical stack

The simplest case of large-area structures, beyond that of the single top-interface of a bulk sample, is that of a stack of planar layers with thicknesses comparable to or smaller than the wavelength of incident light, infinite transverse extent in the in-plane direction, and located in between two semi-infinite half spaces [see Fig. 5(f)]. In this type of structure, all interfaces are planar and parallel, in which case plane wave incidence allows for analytical treatment.[17] Importantly, only a single reflected and transmitted plane wave exists—and inside the optical stack, just a pair of forward and backward propagating plane waves exist in each layer. While obviously neither plane waves nor transversely infinite interfaces actually exist, this is a good enough approximation for most practical cases. Furthermore, other types of incident fields can be expressed as a series or integral expansion of plane wave components.[22]

2. Gratings

Gratings are essentially planar periodic structures where the period is roughly comparable to the wavelength of light (applications of gratings are discussed in Sec. V A). One-dimensional (1D) gratings consist, for example, of parallel ridges (or corrugations) [see Fig. 5(e)], while two-dimensional (2D) gratings constitute, for example, of crossed ridges or a 2D array of nanoparticles [similarly to the nanowire array in Fig. 5(d)]. In theory, gratings have an infinite extent in the in-plane direction, which greatly facilitates their analytical or numerical treatment. When analyzing the properties of periodic structures, it is sufficient to consider just a single unit cell, i.e., a smaller part of the structure with which the complete structure can be constructed via periodic repetition. This is also a good enough approximation for finite structures as long as the impinging field extends over many periods of the grating such that any edge effects due to the finite extent become negligible. The parameters of a grating affecting its function and performance are basically the shape and dimensions of the repeating structures, the periodicity, and the materials involved. Both dielectric and metallic grating structures have been used, depending on the application.

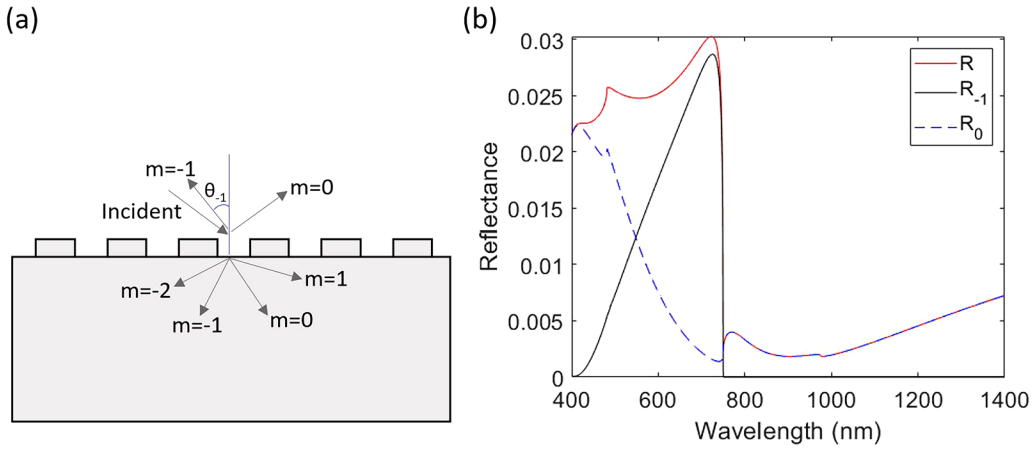

FIG. 8. (a) Schematic of a 1D SiO_2 grating with air on top. (b) Reflectance of the array in (a) for a period of $P = 500$ nm and grating lines with 250 nm width and 350 nm height. The incidence angle is $\theta_{inc} = 30°$ and we consider TM polarization (the incidence plane is perpendicular to the grating lines, that is, we consider $\varphi_{inc} = 0$). The simulations were performed with FMM using the refractive index of SiO_2 from Ref. 32. Here, we show also the intensity carried by the 0th and −1st diffraction orders (here we denote R_{m_x,m_y} with R_{m_x} since $m_y = 0$ in this system that has translational invariance in the y direction), which are the only propagating diffraction orders on the reflection side in this wavelength range. The minor kink around $\lambda = 976$ nm corresponds to the wavelength where the −1st diffraction order in transmission becomes evanescent. In transmission, the +1st and −2nd diffraction orders become evanescent at $\lambda = 482$ and $\lambda = 491$ nm, respectively, and we find kinks also in that wavelength region in the reflectance curves.

Furthermore, with 1D gratings, whether the polarization of an impinging field is parallel or perpendicular to the direction of periodicity makes a significant difference in the interaction.

Diffraction gratings operate based on the diffraction of the impinging light field (i.e., via scattering and interference) and can direct the light to certain modes, the so-called diffraction orders (or Floquet modes), in transmission or reflection (see Fig. 8). Let us denote the refractive index on the incidence side as $n_{\mathrm{inc}}(\lambda)$ and on the transmission side as $n_{\mathrm{tr}}(\lambda)$ (here assumed real-valued, but the discussion and derivation can be generalized to an absorbing substrate on the transmission side). We assume that the array has the period p_x (p_y) along the x (y) direction. Next, we consider light that is incident from a direction given by the angles θ_{inc} and φ_{inc}. Then, the incident k-vector has magnitude $k_{\mathrm{inc}}(\lambda) = k_0 n_{\mathrm{inc}}(\lambda) = \frac{2\pi}{\lambda} n_{\mathrm{inc}}(\lambda)$, x-component $k_{\mathrm{inc},x}(\lambda) = \sin(\theta_{\mathrm{inc}})\cos(\varphi_{\mathrm{inc}})k_{\mathrm{inc}}(\lambda)$, and y-component $k_{\mathrm{inc},y}(\lambda) = \sin(\theta_{\mathrm{inc}})\sin(\varphi_{\mathrm{inc}})k_{\mathrm{inc}}(\lambda)$. By Bloch's theorem, the allowed values for the in-plane k-vector in the system are given by $k_{x,m_x}(\lambda) = k_{\mathrm{inc},x}(\lambda) + 2\pi m_x/p_x$ and $k_{y,m_y}(\lambda) = k_{\mathrm{inc},y}(\lambda) + 2\pi m_x/p_y$, where m_x and m_y are integers. These allowed k-vectors correspond to diffraction orders, that is, plane waves, in the homogenous incidence and transmission side. The propagation direction of the (m_x, m_y) diffraction order can be calculated by first calculating its z-component: $k_{z,m_x,m_y,\mathrm{R(T)}}(\lambda) = ((k_0 n_{\mathrm{inc(tr)}}(\lambda))^2 - (k_{\mathrm{x},m_x}(\lambda))^2 - (k_{y,m_y}(\lambda))^2)^{1/2}$. If $k_{z,m_x,m_y,\mathrm{R(T)}}(\lambda)$ turns out imaginary-valued, we are considering an evanescent diffraction order that decays exponentially away from the grating. In contrast, if $k_{z,m_x,m_y,\mathrm{R(T)}}(\lambda)$ is real-valued, we are considering a diffraction order that carries energy away from the grating region. The angle at which such a diffraction order propagates can be calculated from $\theta_{m_x,m_y,\mathrm{R(T)}} = \arctan(((k_{x,m_x}(\lambda))^2 + (k_{y,m_y}(\lambda))^2)^{1/2}/k_{z,m_x,m_y,\mathrm{R(T)}}(\lambda)])$ and $\varphi_{m_x,m_y,\mathrm{R(T)}} = \arctan2(k_{y,m_y}(\lambda), k_{x,m_x}(\lambda))$. With increasing wavelength, each diffraction order, except the one with $m_x = m_y = 0$, transforms from propagating to evanescent type. The wavelength at which this transformation occurs is obtained by solving for λ from $k_{\mathrm{z},m_x,m_y,\mathrm{R(T)}}(\lambda) = 0$. For the propagating diffraction orders, we can calculate the intensity R_{m_x,m_y} (or T_{m_x,m_y}) carried by each mode, that is, the diffraction efficiencies (and each diffraction order can be separated into a TE and TM component with respect to the propagation plane of the diffraction order).[20]

3. Photonic crystals

Photonic crystals are, in the widest sense, structures with a spatially periodic refractive index (the periodicity can occur in one, two, or three directions), and they tend to operate in the region where the wavelength of light is comparable with this periodicity.[50] For example, a 1D photonic crystal could be a stack of planar layers, a 2D photonic crystal could constitute of parallel pillars (or holes) arranged in a 2D transverse lattice embedded in a medium, and a three-dimensional (3D) photonic crystal could constitute of identical inclusions arranged in a 3D lattice embedded in a medium. Typically, the refractive indices involved are real-valued or have a very small imaginary part at optical frequencies, i.e., the materials exhibit negligible optical losses.

The key phenomenon in photonic crystals is the formation of a photonic band structure and accompanying photonic bandgaps, i.e., frequency bands for which there exist no solutions for propagating modes in the photonic crystal. A photonic bandgap can be thought to essentially arise due to Bragg scattering, i.e., due to destructive interference between field components scattered by the periodic structure (similar to the electronic band structure of solids, as arising from the Schrödinger equation). It is of note that the bandgaps only apply for the directions where the photonic crystal exhibits periodicity, and, therefore, 1D and 2D photonic crystals may still support modes with propagation perpendicular to the lattice direction and plane, respectively. Furthermore, with 2D and 3D photonic crystals, there can also exist directional photonic bandgaps, in which propagating modes are forbidden for some directions and allowed for others (again, only considering directions in the plane of periodicity for 2D photonic crystals). Notably, for an emitter in a 3D photonic bandgap, with its emission frequency within a 3D photonic bandgap of the photonic crystal, inhibited radiative decay occurs due to the lack of available radiative states for the decay.[51]

Another interesting phenomenon called slow light also occurs in photonic crystals and results from the dispersion of the photonic bands. The propagation speed of a light pulse envelope is given by the group velocity (v_g), obtained as the first order derivative of the dispersion ($v_g = d\omega/dk$), which converges to very low values near the edges of the bands at the photonic bandgap. However, the extent and usefulness of this phenomenon tends to be limited by a trade-off with bandwidth and distortion from higher order dispersion. Nevertheless, the slow light effect in photonic crystals could find applications, e.g., in lasers and optical amplifiers due to the enhanced effective interaction length or as buffers in optical computation.[52] For a more detailed introduction to the physics of photonic crystals, including photonic bandgaps and slow light effects, the readers are referred to the various textbooks written on the subject, including that of Ref. 50.

4. Aperiodic structures

The term aperiodic structure is used here to refer to a set of structures that lacks periodicity throughout their extent, i.e., including random or quasi-random, deterministic aperiodic, and piecewise periodic structures. An example of random structures would be an array of semiconductor nanowires grown on a substrate via randomly dispersed Au nanoparticles or a rough randomly etched surface with transverse feature sizes much smaller than the wavelength [like in black-Si;[53] see Fig. 1(b)]. Alternatively, a surface defined through colloidal hole-mask self-assembly lithographic patterning yields typically a structure with short range order and a varying degree of long range order.[54] Deterministic aperiodic structures, on the other hand, include fractal patterns and other ordered but not periodic arrangements.

Random structures exhibit some interesting effects related to light scattering, which can be exploited in certain applications. For example, scattering in sufficiently dense random structures can lead to Anderson localization of the light field.[55] In essence, this localization is the result of closed scattering paths with constructive interference becoming more likely than diffusive scattering out of the structure when the scattering mean free path is sufficiently small (upon reaching this condition, the system makes a transition from

diffusive transport to localized states). Such field localization effects have been exploited in, e.g., random lasers where they provide the feedback mechanism to reach sufficient gain for lasing.[56]

V. EXAMPLES OF APPLICATIONS

To illustrate the prospect of nanophotonics, we give an overview of a few selected applications, as well as some of the specific optical effects arising in these cases. We start with the optics of gratings and anti-reflection coatings, continue with nano-antennas, nano-resonators, and single-photon sources, then move to optoelectronic applications highlighting nanophotonic aspects of LEDs and solar cells, and end with thermophotonics (TPX) for electroluminescent cooling.

A. Gratings and diffractive optics

In the past several decades, diffraction gratings with different periods have found extensive use in diverse areas. In this context, it has become critical to accurately predict the diffraction efficiencies of the propagating diffraction orders of different types of gratings (see Fig. 8 for examples of the diffraction orders). If the grating period is far larger than the wavelength of light, one can usually model the grating as a thin complex-amplitude-modulating screen. In such circumstances, it is possible to employ simple scalar methods consistent with Fourier optics[57] to predict the efficiencies. 2D diffractive gratings can also be designed to operate as various optical elements,[58] such as polarizers and wave-plates.

When the grating period approaches the wavelength of light, the diffraction efficiencies become polarization-dependent, necessitating the use of electromagnetic theory.[59–61] In this so-called resonance domain, only a few orders propagate (others being evanescent), and several new features appear in the efficiency curves, which can only be predicted by an exact solution of the grating-diffraction problem. The most prominent of these features are anomalies caused by, e.g., plasmon excitation in linear metallic gratings in TM polarization and guided-mode excitation in coated dielectric gratings. These effects, which allow efficient coupling of light into guided surface waves or modes of thin film waveguides,[62–64] show up as rapid variations in the efficiency curves of propagating orders when either the illumination wavelength or the angle of incidence is varied (see Fig. 9).

By modulating the period of a grating as a function of position, one can perform rather arbitrary wavefront transformations, such as conversion of a plane wave into a spherical wave (diffractive lens) or shaping a Gaussian beam profile into some more desirable form at a target plane. On the other hand, one can modulate the internal structure of a single grating period to form beam arrays with a desired distribution of intensities among the beams. These are examples of diffractive elements.[65,66] For example, grating couplers are designed to use diffraction to couple modes propagating in a waveguide to freely propagating modes or vice versa.

If all features in the diffraction structure have transverse dimensions substantially larger than the wavelength, simple modeling based on the complex-transmittance approach are available. If not, an exact electromagnetic analysis is required.[67]

B. Subwavelength structures and metamaterials

If the grating period is reduced sufficiently below the wavelength, the grating equation dictates that only the zeroth reflected/

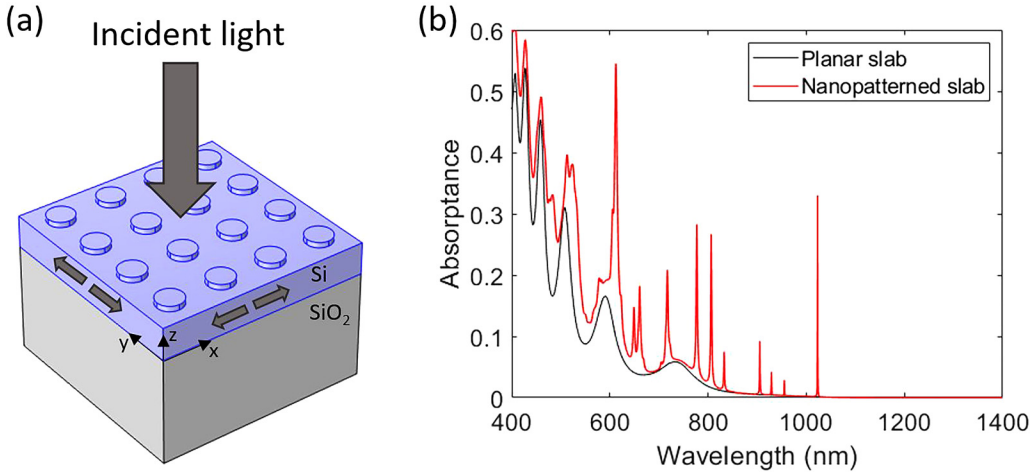

FIG. 9. (a) Schematics of a planar Si slab with a 2D grating of Si nanodisks on top. The arrows in the *x*–*y* plane inside the slab indicate the in-plane waveguide modes whose excitation is enabled by the periodic nanopatterning. (b) Absorptance of normally incident *x*-polarized light. Here, the thickness of the Si slab is 300 nm, the period of the nanodisk array is 400 nm, and the height of the nanodisks is 50 nm. The sharp peaks are assigned to excitation of in-plane waveguide modes, which enhance the absorption. For comparison, we show also the corresponding absorptance of a 300 nm thick unpatterned Si slab, for which we see a smoother absorptance that is modulated by the interference within the slab in the out-of-plane direction (that is, in the *z* direction). The simulations were performed with FMM using a 0.1 nm step in wavelength with the refractive index of Si from Ref. 35 and SiO_2 from Ref. 32 (see also Fig. 4 for the values used).

transmitted orders propagate. In this subwavelength domain, the grating behaves much as a thin film, but the structuring allows for the realization of unusual optical properties.[67] Even though simple theories of effective media and form birefringence are available in special cases, exact electromagnetic theory is generally required for the analysis and design of such subwavelength structures (also known as metasurfaces[68] or, under certain conditions, also as 2D photonics crystals).

Subwavelength structures can be either metallic or dielectric. One example of dielectric structures is an anti-reflection layer consisting of a subwavelength-period array of pillars or pyramids, mimicking a moth's eye.[69] On the other hand, an inductive grid[70] consisting of a subwavelength-period array of holes pierced in a thin metal screen acts as a low-pass filter. Such a structure can, near a resonant wavelength, transmit more light than one would expect on purely geometrical arguments,[71] up to 100% for a perfectly conducting screen.[70]

The concept of subwavelength area coding[72] can be applied to design binary structures that mimic the operation of any diffractive surface-relief elements such as a diffractive lens.[73] These structures (sometimes referred to as metalenses)[74] can be realized using a number of different geometrical shapes, including pillars with circular or rectangular cross section.

In general, metamaterial structures essentially consist of various arrangements (periodic or aperiodic) of nanoparticles or nano-antennas. Among the notable applications of metasurfaces is shaping and redirecting reflected or transmitted light fields in order to either shrink conventional optical components (like lenses) or to provide new functions.[75–78] Metamaterials can be designed to show a strongly anisotropic effective dielectric tensor in a way that gives rise to hyperbolic dispersion for light in the metamaterial.[79] The hyperbolic dispersion can, for example, strongly enhance the spontaneous emission from an emitter within the metamaterial.[79] Overall, metamaterials and their applications are a rather wide research field and a detailed discussion goes beyond the scope here. For the interested readers, we recommend Refs. 75–77 and 79–85.

C. Anti-reflection coatings or surfaces

The reflection of an incident plane wave at an interface between two different materials is one of the most relevant optical effects in many applications. The reflection can be described by the Fresnel equations (see Sec. VI A 1 for details), which are a solution to the Maxwell equations for a system consisting of a single planar interface between materials of $n_1(\lambda)$ and $n_2(\lambda)$. At normal incidence, the reflectance, that is, the fraction of incident intensity that is reflected, is given by $R(\lambda) = |n_1(\lambda) - n_2(\lambda)|^2/|n_1(\lambda) + n_2(\lambda)|^2$. Therefore, for an interface between air (with $n = 1$) and a typical semiconductor like Si with $n \approx 4$ in the visible wavelength range, $R = 36\%$ for normally incident light [see Figs. 10(b) and 10(c)].

Typically, we wish to minimize such reflections, which can be done with anti-reflection coatings. An optimum anti-reflection coating is broadband, omnidirectional, and polarization insensitive and has a low cost and high stability. Anti-reflection coatings are widely used for device and optical component production, for example, in glasses, mirrors, camera lenses, solar cells, photodetectors, LEDs, surface emitting lasers, flat panel displays, and in optical sensing and imaging applications.[86]

The simplest way for reducing reflections is to use interference within a thin film layer to suppress reflections. The thickness of such a film is typically chosen to be $\lambda/(4n_{\mathrm{ARC}})$-thick to give a destructive π phase shift to the reflected light (here, n_{ARC} is the refractive index of the anti-reflection layer at the target wavelength where R should be minimized). The drawback of such a single-layer anti-reflection coating is that it shows a narrow bandwidth and strong incidence angle dependence. With multi-layer anti-reflection coatings, it is possible to increase the bandwidth and reduce the sensitivity to incidence angle. However, finding materials for multi-layer anti-reflection coatings can become demanding, both with regard to available refractive indices and available deposition methods.

Therefore, a promising path for anti-reflection coatings enabled by nanophotonics is to use nanostructured arrays where the filling factor of the high refractive index material increases

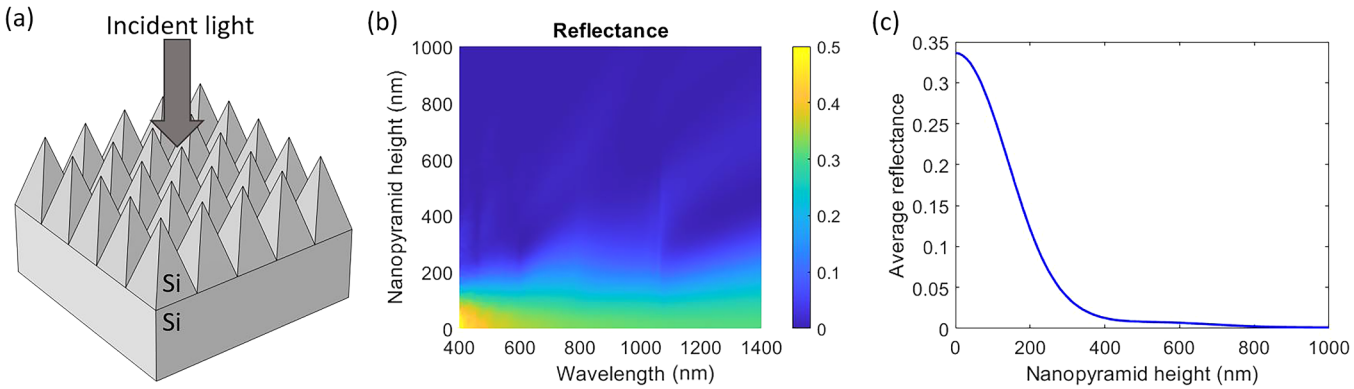

FIG. 10. (a) Schematics of a Si substrate with a Si nanopyramid patterning on the top surface. (b) Reflectance of normally incident light on the nanopyramid array. Here, the period of the nanopyramid array is 300 nm. (c) Wavelength averaging of the results shown in (b). The simulations were performed with FEM using the refractive index of Si from Ref. 35 (see also Fig. 4 for the values used). For a pyramid height of 1000 nm, the average reflectance is <0.2%.

gradually when moving further into the anti-reflection coating from the incidence side (see Fig. 10). Then, we can imagine the anti-reflection coating to consist effectively of infinitely many thin film layers with a gradually increasing refractive index. However, for proper design, a solution for the Maxwell equations for the full 3D optics problem is needed since the anti-reflection coating works in the diffraction optics regime, with a subwavelength unit cell. Such structures, inspired by the moth's eye,[69] exhibit promising broadband and quasi-omnidirectional antireflective properties.[86]

D. Nano-antennas

Nano-antennas can be understood as nanostructures that either redirect energy from a local light field into an outgoing freely propagating light field, i.e., act as transmitters, or redirect energy from an incident freely propagating light field into fields inside local structures, i.e., act as receivers. At radio frequencies, antennas are metallic structures with features proportional to the wavelength of the radiation, and the whole antenna design can be scaled with the wavelength. However, this scaling does not extend to optical frequencies, since field penetration inside the metal, that is, the skin depth, becomes comparable to the size of the nanostructures, which leads to a complicated size-dependence in the optical response. Therefore, nano-antennas employ novel geometries and can consist of metals, dielectrics, semiconductors, or hybrids thereof.[87]

The performance of a transmitting antenna can be assessed by the efficiency with which it extracts power from the local emitter and by the directivity toward a receiver, i.e., the fraction of power that is carried from the emitter by freely propagating modes toward the receiver. Taken together, these two measures define the antenna gain. For example, a semiconductor nanowire nano-antenna with a cylindrical shape and a sharp tip[88,89] can funnel the spontaneous emission from an embedded quantum dot emitter to a guided wave mode along the nanowire and then out-couple this mode via the tip to a freely propagating mode with a Gaussian far-field profile. Furthermore, it has been suggested that the interference between electric and magnetic resonances in a nanoparticle nano-antenna can be exploited to tune the antenna directivity when the light originates from a nearby dipole emitter.[90] A nano-antenna can also greatly enhance the spontaneous emission rate of an emitter, which can usually be attributed to the Purcell effect, thus improving the efficiency by also increasing the radiated power and not just the efficiency with which it is collected. For example, a metallic nano-antenna laid across a rod-shaped piece of semiconductor acting as the emitter was observed to enhance the spontaneous emission rate by an estimated factor of 115.[91] It could also be noted here that the performance of such a transmitting metallic nano-antenna could also be analyzed in the framework of an equivalent electric circuit model.[91]

A receiving nano-antenna, on the other hand, couples the incident light field to strong, localized fields. As mentioned in Sec. IV A, plasmonic nanoparticles and oligomers can effectively transfer energy from an incident field to strong, highly localized electric fields, while dielectric or semiconductor nanoparticles and oligomers are more suited to transferring energy to localized magnetic fields. For example, in Ref. 92, an Au nano-antenna was used to concentrate IR light to a deeply subwavelength active volume of Ge to form a functional photodetector.

E. Resonator cavities

Compared to the nano-antennas in Sec. V D, which optimize coupling between radiation and local light fields, resonator cavities are instead designed to confine light. Nanoscale resonator cavities can be nanostructures with a simple geometry such as a sphere, disk, ring, or more specific as toroidal, tubular (including capillary and bottleneck resonators), or microbubble, designed to confine light to a small spatial volume of subwavelength in extent (see, for example, Ref. 93).

The, arguably, simplest cavity structure is a Fabry–Pérot resonator consisting of two highly reflecting mirrors that are placed parallel to each other at some distance (see Fig. 11). The light travels in the cavity and reflects from the mirrors, producing a resonant standing optical wave. The constructive and destructive interference of the reflected waves determine the transmission maxima and minima that are affected by the optical path length within the cavity (see Fig. 11). The cavity performance is determined by the following parameters: the finesse F (proportional to the number of interfering beams within the cavity) and the quality factor Q (the ratio of the stored energy to the energy dissipated per radian of the oscillation). Both parameters can be improved by changing the reflectivity of the mirrors. Additionally, the Q factor can be modified by changing the cavity length. The resonator has found many applications, particularly in lasers where one of the mirrors is characterized by a slightly higher transmission in order to guide radiation into one specific direction. High-Q laser resonators can be used for obtaining laser output with a very narrow linewidth.

High-Q dielectric resonators, like microrings and photonic crystal micro-cavities, are enabling the use of microphotonic circuits for telecommunications and fundamental research, such as filtering, sensing, nonlinear interaction, lasers, quantum optics, and metrology.[94] In Ref. 95, optical cavities have been summarized and categorized into two groups: standing-wave cavities and traveling-wave cavities. The photonic crystal cavity and distributed feedback cavity are considered standing-wave cavities while the ring cavity is a typical example of a traveling-wave cavity. In Ref. 96, a review is given of optical resonators based on whispering gallery modes, which are a type of traveling waves. This type of resonators is typically characterized with a high Q factor and, consequently, has low losses.

F. Single-photon sources

The controlled emission of single photons can be applied for many applications, such as quantum communication, cryptography, computation, detecting, sensing, non-classical light sources, quantum metrology, and accurate measurements.[97,98] A single-photon source is expected to be mechanically robust, have single-photon purity, indistinguishability, high efficiency, high brightness (that is, high maximum single-photon emission rate), ease of fabrication, and cover through proper design as wide of a wavelength range as possible.[98]

Nowadays, there are many available quantum systems that are suitable for single-photon emitters, although the most promising is typically considered to be a solid-state single-photon source based on an artificial atom embedded into a system that geometrically will support and control the propagation, direction,

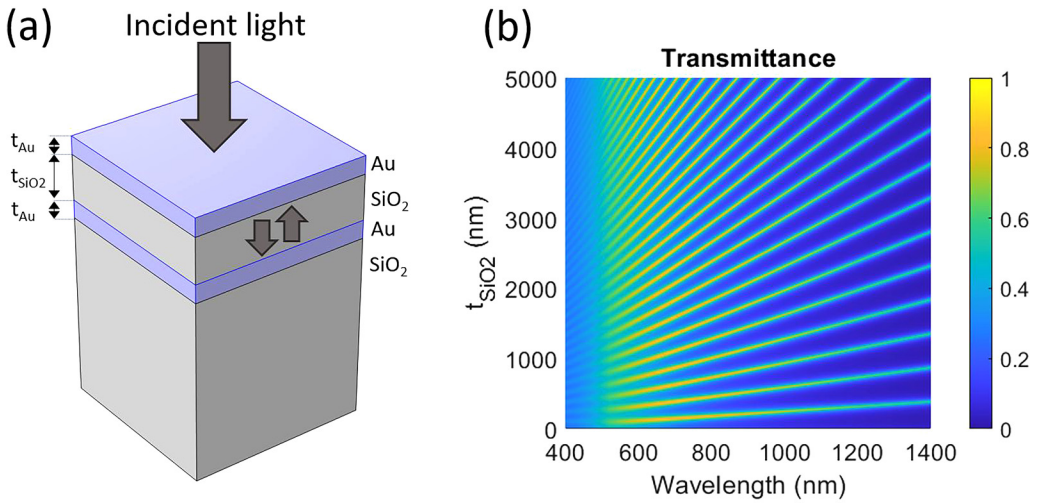

FIG. 11. (a) Schematics of a Fabry–Pérot cavity consisting of a SiO_2 layer of thickness t_{SiO_2} between two Au mirrors of thickness t_{Au} on top of a SiO_2 substrate. The arrows inside the cavity indicate the counter-propagating plane waves inside the cavity. (b) The transmittance of normally incident light through the cavity in (a) for t_{Au} = 10 nm, simulated with FMM using the refractive index of Au from Ref. 34 and SiO_2 from Ref. 32 (see also Fig. 4 for the refractive index values used).

and polarization of the light.[97,98] For such control, for example, micropillar cavities, circular dielectric gratings, photonic crystal cavities, and microlenses have been considered.[99] As an example from nanophotonics, semiconductor nanowires with an embedded quantum dot single-photon emitter show great potential for single-photon source applications due to the nanowires' inherent waveguide properties, leading to efficient light coupling, high directivity, and polarizability of the emitted light.[98] Optimization of such properties defines a challenging design problem in nanophotonics modeling.[100] In such modeling, the single-photon emitter can typically be considered a classical dipole emitter, which captures the most important optical aspects, such as the Purcell effect, coupling of the single-photon emitter into varying optical modes, and emission directionality.[98]

G. LEDs

LEDs are driven by electroluminescence, where excess electrons and holes are supplied to the active region of a semiconductor (a p–n junction or, more commonly, a quantum well) by an external electric source. The excess electrons and holes subsequently recombine, leading to the emission of photons. As compared to the spatially localized point-like emitter in single-photon sources in Sec. V F, in an LED, the emitter region is typically of larger spatial extent (area) to allow for larger output intensity. The range of possible applications for LEDs is broad, e.g., general illumination, different types of displays, short-haul communications, optoelectronic computer interconnects, and even solid-state cooling.[101,102]

For accurate electro-optical modeling of LEDs, accounting for coupled electron–hole transport should be performed to include effects, for example, from (i) the injection of charge carriers to the active region to induce net radiative recombination and (ii) non-radiative recombination through the Shockley–Read–Hall (SRH) recombination in the active region, at sidewalls, and at other interfaces, and (iii) Auger recombination at high injection level,[29,103,104] in addition to radiative recombination processes. However, much of the LED design can be performed with purely optical modeling, which predicts the Purcell factor, the extraction efficiency of internally emitted photons, the amount of parasitic absorption of emitted photons, and emission directionality and polarization to the free space outside the LED.[29] Reduced non-radiative recombination, increase of radiative recombination rate, and increased extraction efficiency are the guiding direction for higher performance LEDs.[105] The external quantum efficiency (EQE) of an LED becomes an important optimization parameter. The EQE can be expressed as a product of the IQE and the extraction efficiency.[106] The EQE shows the probability that an injected electrical charge carrier results in the emission of photons to the exterior of the LED. In the typical case where a recombination event can give rise to maximally one emitted photon, EQE < 100%, while for some tailored systems, EQE > 100% is predicted.[107]

In Ref. 108, the influence of adopting anti-reflection coatings for LEDs has been highlighted. However, even with a perfect anti-reflection coating, the high refractive index of the semiconductor in the LED active region limits severely the size of the escape cone through which internally emitted photons can be extracted from a planar LED. For a typical semiconductor, only 4% of the internally emitted photons are within the escape cone, the rest experiencing total internal reflection. The potential of increasing the LED extraction efficiency by using photonic crystals was highlighted in Ref. 108. Also, dielectric gratings have been used to improve the light extraction from LEDs.[109] In both cases, the nanostructures circumvent, thanks to diffraction effects, the limited escape cone of the corresponding planar structure. On the other hand, metallic gratings can significantly enhance the emission rate (via the Purcell effect) for emitters close enough to the grating by

allowing the emitter to couple to surface plasmon modes.[110] At the same time, the grating structure can efficiently scatter light from the surface plasmon modes to freely propagating modes outside the LED, increasing the extraction efficiency. Unfortunately, parasitic absorption losses introduced by the metal tend to limit the overall LED efficiency enhancement attainable with such metallic gratings.[111]

Thus, nanophotonics offers for LEDs the prospect of increased IQE (through the Purcell effect), enhanced extraction of emitted photons, and built-in shaping of the output directionality and polarization without the need for bulky external optical components.

H. Solar cells

Solar cells provided by the end of year 2019 almost 3% of the world electricity demand, with an increase in installed capacity by an impressive factor of 30 since the year 2009.[112] This clean source of energy is predicted to continue its growth in the coming years, and any improvements on solar cell design and performance are expected to have a large impact on the global society.

A solar cell can be considered the reciprocal of an LED. In a solar cell, incident light creates excess charge carriers that are extracted from the device as an electric current, in contrast to the LED where excess carriers are electrically injected to the device to result in emission of light out from the device. Currently, a record efficiency of 47.1% has been shown in a solar cell consisting of six III-V compound semiconductor subcells stacked on top of each other in a tandem configuration.[113] The market-dominating single-junction Si solar cell technology has demonstrated laboratory cells with an efficiency above 26%.[114] Nanophotonics offers prospects to enhance solar cells, for example, through (1) nanostructured anti-reflection coatings,[86] (2) nanophotonic light trapping to reduce the absorber thickness,[115] and (3) nanostructured absorber region to open up for new materials configurations.[116]

By modifying the surface of a solar cell to enhance total internal reflection of the entering light, the optical path length of the light inside the active layer can be increased, and consequently the absorption is enhanced.[117] In a ray optics description, such light-trapping scheme gives an absorption enhancement by up to a factor of $4n^2$ for an active region of a refractive index n.[115,117] In the nanophotonics regime, varying geometry arrangements show promise for beating this conventional light-trapping limit, including, e.g., 1D periodic semiconductor or dielectric structures, 2D diffraction gratings and photonic crystals, plasmonic structures, nanowires, and randomly structured semiconductor surfaces.[115,118,119]

As an example of using nanostructuring for the whole absorber region, in 2009, it was shown that a single semiconductor nanowire could be used to strongly alter the absorption of light.[120] By designing the diameter of the horizontal nanowire, analytically predicted absorption resonances through leaky-mode resonances were also found in measurements.[120]

Such single-nanowire absorption resonances are promising for solar cells and photodetectors.[120–122] In parallel, for large-area solar cells and photodetectors, nanowire arrays gained interest.[4,89,123–128,129–134] For example, already in 2013, a semiconductor nanowire array solar cell, consisting of 4×10^6 vertical nanowires connected in parallel [see Fig. 1(d)], demonstrated an efficiency of 13.8%,[4] thanks to enhanced absorption through guided modes in the nanowires.[126,135]

I. Electroluminescent cooling and thermophotonics

Contrary to widespread belief, the application of LEDs is not limited to solid-state lighting. In fact, LEDs are thermodynamic machines that can also be used for solid-state cooling[102] when an LED breaks the unity wall-plug efficiency (WPE) barrier.[136] LED refrigeration is enabled by a phenomenon referred to as electroluminescent cooling,[136] which is experienced by an electrically biased LED when the extracted optical power is higher than the input electrical power, corresponding to an above-unity electricity-to-light conversion efficiency (i.e., WPE > 1). The cooling originates from the capability of semiconductor materials, such as GaAs, to maintain an electronic excitation that enhances the spontaneous emission beyond its thermal value. In a semiconductor LED, with an excitation energy qU_b (U_b being the LED bias voltage), photons with an average energy $\approx E_{bg} + k_B T$ are emitted, even if $qU_b < E_{bg}$. The remaining energy is drawn from lattice heat (thermal energy $\approx k_B T$, with k_B the Boltzmann constant and T the temperature), which thus has the prospect of cooling.[102] Then, $\mathrm{WPE} = \mathrm{EQE} \times (E_{bg} + k_B T)/qU_b$ can exceed unity (the cooling threshold) for an appropriate combination of the EQE (<100%) and bias ($qU_b < E_{bg}$). At such a condition, the amount of emitted optical energy exceeds the parasitic heating of the LED, which is caused, e.g., by the non-radiative relaxation processes, leading to cooling.

To fully harvest the potential of electroluminescent cooling, a highly efficient LED can be included in a thermophotonic configuration,[136] where the emitted optical energy is recycled by an absorbing photovoltaic device (we refrain from using the term solar cell here since it is not sunlight incident on the device). Thermophotonics (TPX) is the concept of harnessing the inseparable thermodynamics of electroluminescence (superthermal emission) from LEDs and photovoltaic energy production. For electroluminescent cooling, TPX is expected to boost the coefficient of performance to levels that are out of reach for other competing or established technologies, such as photoluminescent cooling or thermoelectricity.[102] TPX provides an environmentally friendly framework for cooling and energy production, as it uses photons as the heat carriers to achieve solid-state cooling (without employing harmful refrigerant substances) or generate electrical power from waste heat, offering huge and new opportunities for harvesting thermal energy from the abundant waste heat sources, outside the original concept of solar energy harvesting.[137] In these devices, nanophotonics could play a crucial role, for example, in the design and implementation of a vacuum nanogap layer between the LED and the photovoltaic device. Such a layer would minimize conventional lattice heat transfer (which counteracts the TPX operation) but allows for efficient optical near-field heat energy transfer from the LED to the photovoltaic device.[102,137]

VI. BRIEF DESCRIPTION OF SELECTED SIMULATION METHODS

A. Analytical methods

The strength of analytical methods lies in their ability to give exact results for the light scattering problem (in the form of exact

solutions to the Maxwell equations). Therefore, analytical solutions are excellent for validating the accuracy of numerical solutions. Furthermore, analytical methods can give insights into the underlying physics of the optical response, which can at times be difficult to extract from fully numerical methods. However, analytical solutions exist only for a limited set of special (highly symmetric) geometries.

1. Single planar interface and thin film

For a plane wave incident on a planar interface at an angle θ_i from the surface normal, the reflection and transmission of the plane wave are also of plane wave form and given by well-known formulas. The angle of reflection is simply equal to the angle of incidence ($\theta_r = \theta_i$), while the angle of transmission (θ_t) is given by Snell's law,

$$n_i \sin(\theta_i) = n_t \sin(\theta_t), \tag{6}$$

where n_i is the refractive index of the incident side medium and n_t is the refractive index of the transmission side medium. The reflection and transmission coefficients, which relate the electric field of the reflected and transmitted plane wave to that of the incident plane wave, are given by the Fresnel formulas and depend on the polarization of the incident light (here, for simplicity, we assume materials with real-valued refractive index; for the more general case of a medium on the transmission side with complex-valued refractive index, the readers are referred, for example, to Ref. 138),

$$r_s = \frac{n_i \cos(\theta_i) - n_t \cos(\theta_t)}{n_i \cos(\theta_i) + n_t \cos(\theta_t)}, \tag{7}$$

$$t_s = \frac{2n_i \cos(\theta_i)}{n_i \cos(\theta_i) + n_t \cos(\theta_t)}, \tag{8}$$

$$r_p = \frac{n_i \cos(\theta_t) - n_t \cos(\theta_i)}{n_i \cos(\theta_t) + n_t \cos(\theta_i)}, \tag{9}$$

$$t_p = \frac{2n_i \cos(\theta_i)}{n_i \cos(\theta_t) + n_t \cos(\theta_i)}, \tag{10}$$

where r_s is the reflection coefficient for s-polarization, t_s is the transmission coefficient for s-polarization, r_p is the reflection coefficient for p-polarization, and t_p is the transmission coefficient for p-polarization. The reflectance is given by $R = |r|^2$ and for a non-absorbing transmission side medium, the transmittance, that is, the fraction of incident intensity that is transmitted, is given by $T = \frac{n_t \cos(\theta_t)}{n_i \cos(\theta_i)} |t|^2$ (see Ref. 139 for generalization to absorbing transmission side medium).

In the case of a single layer, there are two planar interfaces at which reflection and transmission of the field takes place. In addition to the initial reflection and transmission at the first interface, the field inside the thin film, which is in the form of a plane wave, experiences multiple (that is, infinitely many) reflections inside the layer (with partial transmission out of the layer, with each transmission contribution interfering with the previous contributions) at the interfaces while phase delay accumulates during propagation between the interfaces. As a result, summing the infinitely many field components yields a geometric series, and the resulting reflection and transmission coefficient can actually be obtained analytically [see, e.g., Eqs. (7.1)–(8) in Ref. 17 for the resulting equations and Eq. (S15) in supplementary information of Ref. 64 for an example of such a derivation generalized for a single-layer grating).

2. Mie theory for spheres

For the thin film above, the translational invariance in the in-plane direction guarantees plane wave solutions. Another system of high symmetry is the sphere, for which the so-called Mie theory, which in reality is a mathematical calculation method rather than a theory, supplies solutions for the Maxwell equations in terms of spherical harmonics.[41,140] For an exact solution, infinitely many such harmonics should be used, but in practice the expansion is truncated at some large number of harmonics to yield sufficient accuracy in the results. The Mie theory yields the full solution to the Maxwell equations in terms of the spatially resolved electric field around the particle as well as the polarization and angle dependent scattering to the far-field[41,140] and applies also for cylinders.[41] Note that the well-known Rayleigh scattering,[141] which explains why small dielectric particles scatter shorter wavelengths more efficiently than longer wavelengths, follows from the Mie theory when $D \ll \lambda$ (that is, the small-diameter or alternatively long-wavelength limit), where electrostatics applies for the analysis. Spherical harmonics are used also in the numerical Waterman's T-matrix approach for non-spherical particles and composites of particles.[142]

B. Other analytical tools

1. Effective medium theory

Whereas the above analytical treatment of the thin film and the Mie theory for the sphere yield solutions to the Maxwell equations, effective medium theory is a simplification to help the analysis of optical response from nanostructures made of a multiphase medium consisting of underlying materials mixed at the (deep) subwavelength scale. Effective medium theory provides a description of the optical response of the multiphase medium starting from the refractive indices and the fractions of the constituent materials forming the multiphase material. The effective medium theory results in an effective refractive index $n_{\text{eff}}(\lambda)$ for the multiphase material, which can be used in the Maxwell equations for the modeling of light scattering from the nanostructure.

Typically, in effective medium theories, retardation effects are not taken into account, and the optical response is hence calculated in the electrostatic limit [by assuming that inclusions of one material into another surrounding host material are at the (deep) subwavelength scale and randomly distributed]. This approximation allows for simple analytical equations for calculating $n_{\text{eff}}(\lambda)$.[143] Depending on the assumptions made for the type of inclusion, different effective medium theories like the Bruggeman and the Maxwell–Garnett effective medium theory have been used.[143] Indeed, we use implicitly an effective medium theory constantly when using the $n(\mathbf{r},\lambda)$ of bulk materials in the Maxwell equations: $n(\mathbf{r},\lambda)$ describes the effective material response from the underlying

atomic nuclei and electrons at position $\mathbf{r}$ and wavelength λ (see section 6.6 in Ref. 18).

Importantly, we could perform a full solution of the Maxwell equations including the spatial distribution of the constituent materials. However, such a solution would require a very fine spatial resolution in the simulations, making the solution process numerically heavy and possibly prohibitive, beyond available computational resources. Also, if the inclusions are randomly distributed, we probably need ensemble averaging of modeled results from randomly distributed inclusions. The use of the effective medium theory circumvents both these issues and allows us to focus on the resulting diffraction effects from larger geometrical features of the nanostructures instead. However, we must be aware that we use the effective medium theory for the underlying materials response and should assess possible impact of it on the resulting optical response (for example, by using several different effective medium theories to assess how the choice of the theory affects the resulting optical response of the nanostructures).

2. Lorentz reciprocity

Lorentz reciprocity can be used for relating results from the scattering of an incident plane wave to the resulting emission from a dipole emitter within the system. For example, with the reciprocity, we can analyze the far-field emission from nanostructures through scattering modeling,[26] which can, depending on the problem type, be more convenient to perform than actual dipole emission modeling.

Let us consider the case of two dipole emitters, denoted as 1 and 2, located at positions $\mathbf{r}_1$ and $\mathbf{r}_2$ with dipole moments $\mathbf{p}_1$ and $\mathbf{p}_2$. The electric field emitted by dipole 1 and 2 are denoted as $\mathbf{E}_1(\mathbf{r})$ and $\mathbf{E}_2(\mathbf{r})$, respectively. The Lorentz reciprocity, which originates from the Maxwell equations,[144] states that $\mathbf{E}_1(\mathbf{r}_2)\cdot\mathbf{p}_2 = \mathbf{E}_2(\mathbf{r}_1)\cdot\mathbf{p}_1$. Thus, Lorentz reciprocity can be used for analyzing how two emitters couple to each other.

Importantly for the current nanophotonics modeling, by moving emitter 2 toward infinity, the field emitted from dipole 2 toward dipole 1 or from dipole 1 toward dipole 2 resembles a plane wave due to the large distance between the dipoles. Hence, we can use this reciprocity to relate the far-field emission from dipole 1 to a given emission angle (in a certain polarization state) to the local electric field $\mathbf{E}_2(\mathbf{r}_1)$ caused by a plane wave incident from that direction toward dipole 1 in the same polarization state (see supporting information of Ref. 29).

Thus, with the Lorentz reciprocity, we can link scattering and emission modeling, and it applies for any method that can solve the Maxwell equations. For example, the electric field induced by an incident plane wave in and around the nanostructure allows then to quantitatively assess how strongly a dipole emits into that direction for varying dipole position. Similarly, the emission from a dipole allows then to quantitatively assess how strongly an incident plane wave couples to the position of the dipole for varying incidence angle and polarization. Note that due to the three orthogonal dipole moment orientations and the two orthogonal polarization states of a plane wave, either three separate emission simulations or two separate scattering simulations are needed to fully employ the Lorentz reciprocity.

3. Near-to-far-field transformation

When modeling dipole emission with methods like the finite element method (FEM) or finite-difference time-domain (FDTD) method that yield the solution $\mathbf{E}(\mathbf{r},\lambda)$ of the electric field from the dipole emitter in a finite simulation domain around the dipole emitter, a near-to-far-field transformation (NFFT) is needed for extracting information about the emission of light to far away from the nanostructure. In the NFFT, the near-field solution $\mathbf{E}(\mathbf{r},\lambda)$ in the vicinity of the nanostructure is propagated analytically toward infinity, where it yields information of the far-field emission directionality and polarization. An extended solution for the NFFT has been presented in Ref. 145, which can take into account a semi-infinite substrate together with a stack of uniform layers (including the case of no layers). Then, it is possible to obtain the far-field pattern to the bottom (substrate) and top side, as well as the excitation of possible guided modes along the layers (when one or more layers are present).[145]

Note that the Purcell factor can be obtained directly from the near-field solution $\mathbf{E}(\mathbf{r},\lambda)$, as described in Sec. II B, by integrating the Poynting vector over a closed surface around the dipole (and dividing by the emitted power from the same dipole but in a homogeneous medium). Also, as discussed in Sec. VI B 2, we can circumvent the need of the NFFT by obtaining the far-field emission characteristics from the corresponding scattering problem with the Lorentz reciprocity. However, the Lorentz reciprocity does not yield information about parasitic absorption of the emitted light, which is readily available from the near-field solution (by integrating the spatially resolved absorption in the parasitically absorbing region) when modeling with the dipole emitter in the system.

C. Numerical methods

When analytical methods are not available, for example, due to complex geometry, we turn to numerical methods to obtain an approximate solution to the Maxwell equations. Below, we discuss some of the most popular methods employed in nanophotonics and give brief details of their most important technical aspects. The readers are referred to the cited references for additional technical details and numerical implementation of the methods.

1. Transfer matrix method (TMM)

As discussed in Sec. VI A 1, the single thin film layer allows for a closed-form analytical solution for the scattering of a plane wave. With an increasing number of layers stacked on top of each other, a similar closed-form solution becomes intractable. Also with increasing number of layers, due to the translational invariance in the x–y plane, only a single plane wave (that can propagate forward and backward) shows up in each layer.

The transfer matrix method (TMM) allows for a systematic analysis of an optical stack consisting of an arbitrary number of layers (see p. 426–428 in Ref. 146). For illustration and discussion purposes, we outline here a TMM implementation for the case of non-absorbing materials and propagating waves in each layer (for the more general case of decaying waves, due to absorption or total internal reflection, within a system, the readers are referred, for example, to Ref. 139).

Let us consider a system consisting of N layers where layer j has thickness d_j and refractive index n_j. We denote the incidence side as the left side and the transmission side as the right side, using subscripts L and R. On the incidence side, which we could denote also as layer 0, we have refractive index n_L and incidence angle θ_L. The right side, which we could denote also as layer $N+1$, has refractive index n_R.

For each layer, we consider the forward propagating plane wave E_j^+ and the backward propagating plane wave E_j^-. The electric field solution within layer j can be obtained from the solution at the left end of the layer, that is, at $z = z_j$, by propagating the fields by the distance $z - z_j$. Such propagation can be expressed in a matrix form as

$$\begin{bmatrix} E_j^+(z) \\ E_j^-(z) \end{bmatrix} = \mathbf{P}_j(z) \begin{bmatrix} E_j^+(z_j) \\ E_j^-(z_j) \end{bmatrix}, \tag{11}$$

where

$$\mathbf{P}_j(z) = \begin{bmatrix} \exp(ik_{z,j}(z - z_j)) & 0 \\ 0 & \exp(-ik_{z,j}(z - z_j)) \end{bmatrix}, \tag{12}$$

with $k_{z,j} = k_j \cos(\theta_j)$ with $k_j = 2\pi n_j/\lambda$. The angles within the layers satisfy the relationship $n_L \sin(\theta_L) = n_j \sin(\theta_j)$.

The electric fields at the interface between layer j and $j+1$ are related through

$$\begin{bmatrix} E_{j+1}^+(z_{j+1}) \\ E_{j+1}^-(z_{j+1}) \end{bmatrix} = \boldsymbol{D}_{j,j+1} \begin{bmatrix} E_j^+(z_{j+1}) \\ E_j^-(z_{j+1}) \end{bmatrix}, \tag{13}$$

where

$$\boldsymbol{D}_{j,j+1} = \begin{bmatrix} (t_{j,j+1}t_{j+1,j} - r_{j+1,j}r_{j,j+1})/t_{j+1,j} & r_{j+1,j}/t_{j+1,j} \\ -r_{j,j+1}/t_{j+1,j} & 1/t_{j+1,j} \end{bmatrix}. \tag{14}$$

Here, $t_{j,j+1}$ ($r_{j,j+1}$) is the Fresnel coefficient for transmission (reflection) for the plane wave incident from layer j toward layer $j+1$. We solve for either s or p polarized incident light [note that we use in the calculation of these coefficients the angles obtained from the relationship $n_L \sin(\theta_L) = n_j \sin(\theta_j)$].

We can then propagate through all the layers of the system, via matrix multiplications, giving the transfer matrix $\mathbf{M}_{\text{tot}}$ for the system,

$$\mathbf{M}_{\text{tot}} = \mathbf{D}_{L,1}\mathbf{P}_1(d_1)\mathbf{D}_{1,2} \cdots \mathbf{D}_{j-1,j}\mathbf{P}_j(d_j)\mathbf{D}_{j,j+1} \cdots \mathbf{D}_{N-1,N}\mathbf{P}_N(d_N)\mathbf{D}_{N,R}, \tag{15}$$

where element-wise,

$$\mathbf{M}_{\text{tot}} = \begin{bmatrix} M_{\text{tot},11} & M_{\text{tot},12} \\ M_{\text{tot},21} & M_{\text{tot},22} \end{bmatrix}. \tag{16}$$

This transfer matrix relates the field at the right edge of the incidence side to the fields at the left edge of the transmission side by

$$\begin{bmatrix} E_R^+ \\ E_R^- \end{bmatrix} = \mathbf{M}_{\text{tot}} \begin{bmatrix} E_L^+ \\ E_L^- \end{bmatrix}. \tag{17}$$

From the assumption that light is incident from the left, so that $E_R^- = 0$, we obtain

$$E_{\text{L}}^- = -\frac{M_{\text{tot},21}}{M_{\text{tot},22}} E_{\text{L}}^+, \tag{18}$$

$$E_{\text{R}}^+ = \left(M_{\text{tot},11} - \frac{M_{\text{tot},12}M_{\text{tot},21}}{M_{\text{tot},22}} \right) E_{\text{L}}^+, \tag{19}$$

leading to the relationships $r = -\frac{M_{\text{tot},21}}{M_{\text{tot},22}}$ and $t = M_{\text{tot},11} - \frac{M_{\text{tot},12}M_{\text{tot},21}}{M_{\text{tot},22}}$ with $R = |r|^2$ and $T = \frac{n_R \cos(\theta_R)}{n_L \cos(\theta_L)}|t|^2$ (for the generalization to the case of an absorbing substrate such that n_R is complex-valued, please see, for example, Ref. 139).

2. Fourier modal method (FMM)

The Fourier modal method (FMM), also known as the rigorous coupled-wave analysis (RCWA) method, can be seen as an extension of the TMM to a case where the structure shows a periodic variation in one or both of the in-plane directions within a layer. With such periodicity, the optics is no longer described by a forward and a backward propagating plane wave in each layer as in the TMM. Instead, in each layer, infinitely many optical eigenmodes show up, and *modal* in FMM refers to the use of these eigenmodes as a basis for expanding the total electric field.[20] Each of these eigenmodes has a characteristic propagation constant k_j^α for the direction through the layer. Here, j denotes the layer, similarly as in the TMM description above, and α denotes the αth mode in that layer. These modes propagate with an $\exp(\pm ik_j^\alpha z)$ dependence through the layer, similar to the plane wave in TMM. *Fourier* in FMM refers to the use of an underlying Fourier basis, that is, plane waves, in the x–y plane onto which the Maxwell equations are projected, which gives a matrix for each layer. After such a projection, the eigenmodes in each layer are solved for from the corresponding matrix (we solve for all the eigenmodes of the matrix). Hence, the electromagnetic field $\boldsymbol{E}_j^\alpha(x, y)$ [and accompanying $\boldsymbol{H}_j^\alpha(x, y)$] of each eigenmode is expressed in terms of the underlying plane wave basis. The number of considered eigenmodes is determined by the number of plane waves used in the underlying expansion. In other words, by choosing the truncation of that underlying plane wave basis, we fix the size of the finite matrices, which fixes the number of eigenmodes solved for and which in turn affects the accuracy of the simulation. [Here we would like to note the resemblance of FMM with the method of lines (MOL) approach[147] where, similarly, eigenmodes propagating along the z direction are solved for and used for the solution of the full scattering problem. However, in the MOL, instead of projecting the Maxwell equations onto a plane wave basis, a finite-difference scheme is used for numerically approximating the derivatives in the two other directions.]

Thus, in each layer, $\boldsymbol{C}_j^+$ denotes the expansion coefficients of the forward propagating modes and $\boldsymbol{C}_j^-$ denotes the backward propagating modes.[20] By considering the continuity of E_x, E_y, H_x, and H_y at the interface between two layers (which fundamentally follows from the source-free Maxwell equations), we link together the expansions coefficients in adjacent layers (similarly as with the Fresnel coefficients for the single plane wave in the TMM with the $\boldsymbol{D}_{j,j+1}$, but in FMM, we generalize for the multiple eigenmodes present—see Eq. (20) in Ref. 20). Then, we could, in principle, use a transfer matrix approach to connect the incident side to the transmission side [similarly as in Eqs. (15)–(17)]. However, in FMM, since we use an infinite set of eigenmodes (limited only by the numerical truncation), the higher order modes become increasingly evanescent, causing numerical issues due to overflow in floating point precision when propagating the $\exp(-ik_j^\alpha z)$ terms through the respective layer [since we define $\mathrm{Im}(k_j^\alpha) > 0$ for the eigenmodes]. To circumvent such overflow, a scattering matrix approach, where only terms with $\exp(ik_j^\alpha z)$ show up in the actual numerical evaluations in the implementation,[20,148–150] can be employed [see, for example, the notes after Eq. (25) in Ref. 20 for implementing the numerical evaluations]. The use of such a scattering matrix approach can make also the TMM numerically stable in a case when exponentially decaying plane waves show up within the system (either due to absorption or due to total internal reflection).

Note that FMM reduces in essence to TMM if the in-plane corrugation in the system disappears. Formally, there is still an infinite number of eigenmodes in FMM, now in the form of plane waves, in each layer. However, the incident plane wave excites only one eigenmode in each layer—the same single plane wave as in TMM—and we can truncate the underlying plane wave basis in FMM to that single plane wave without loss of accuracy.

In the FMM formulation, the refractive index in each of the layers is assumed to vary only in the x and y directions [and the z-invariance within each layer allows for the $\exp(\pm ik_j^\alpha z)$ dependence in the eigenmodes, which is crucial for the method]. With the staircase approximation,[151] FMM can be extended to cases with slanted sidewalls of the nanostructure, like nanocones.[21] Furthermore, in FMM, it is possible to include a dipole emitter by inserting an artificial interface at the location of the dipole (thus splitting the original layer into two new layers). At that added interface, the continuity of E_x, E_y, H_x, and H_y between the two new layers is modified in a way that corresponds to including the dipole into the system.[149]

3. Finite element method (FEM)

In the finite element method (FEM), the simulation domain is divided, using a meshing process, into finite elements, which are typically triangles in 2D simulations and tetrahedral volumes in 3D simulations. In the simulation, the values for $\mathbf{E}(\mathbf{r})$ (in an electric field formulation of the method) are obtained at the discrete nodes, that is, vertices, of each element.[21,152] Within each element, typically, a second order polynomial is assumed as the shape function for the variation of $\mathbf{E}(\mathbf{r})$, and the values for the polynomial are set by the values at the nodes.[21] Thus, FEM yields as solution the spatial variation of the electric field throughout the simulation domain. The magnetic field can then be evaluated from the electric field using the Maxwell equations.

In the actual FEM simulation, the Maxwell equations are projected onto these finite elements with the help of the shape function. This results in a linear equation set that can be written in the matrix form $\mathbf{Ax} = \mathbf{b}$. Here, $\mathbf{x}$ contains information about the electric field values at the nodes of all finite elements, $\mathbf{b}$ contains information about the boundary conditions, and $\mathbf{A}$ contains information about the geometry and the materials forming the structure. Here, we would like to note that a finite-difference frequency-domain (FDFD) formulation[153] gives rise to a similar equation system, highlighting the large difference to the FDTD formulation (see below) where such linear equation sets do not show up.

In FMM and TMM, the outgoing radiation condition is fulfilled naturally by semi-infinite superstrate and substrate regions into which light can propagate (in the form of plane waves if homogenous regions; otherwise in the form of more complicated eigenmodes if 1D or 2D patterned superstrate or substrate regions are considered in FMM) without backreflection. In FEM, a finite simulation domain is considered, and unphysical backreflection of outgoing radiation must be suppressed to avoid artifacts in the simulation. Typically, perfectly matched layers (PMLs) are used to suppress such backreflection artifacts. PMLs are artificial domains that absorb incident waves, in practice giving the same result as radiation into a domain extending to infinity.[154]

For emission modeling, a dipole can be included directly within the simulation domain in FEM. Alternatively, the Lorentz reciprocity (as discussed in Sec. VI B 2) can be readily used since the $\mathbf{E}(\mathbf{r})$ at the location of the emitter, as induced by the incident plane wave, is obtained directly from the solution process.[26]

4. Finite-difference time-domain (FDTD) method

As the name indicates, the finite-difference time-domain (FDTD) method works in the time domain,[155] whereas the above methods typically work in the wavelength domain (that is, the frequency domain). To solve the Maxwell equations for materials with a response described by $n(\lambda)$, a fitting of the refractive indices into an analytical oscillator form, suitable for time-domain simulations, is needed.[21] In FDTD, the spatial derivatives in the Maxwell equations are approximated with finite differences, and time is propagated forward, typically with a leapfrog time-stepping scheme.[156]

In a typical FDTD scattering simulation, an incident pulse is propagated toward the structure. To obtain wavelength-dependent information, the fields are recorded at selected spatial positions during the time-stepping and Fourier transformed to the wavelength domain afterward. For example, to model the transmission through a grating, the electromagnetic field through the x–y plane just below the grating is recorded for each time step. Then, after the propagation of the pulse is finished (such that a negligibly small field strength exists in the simulation domain), the Fourier transformation yields information of the wavelength-dependent transmitted power. To obtain the transmittance, we normalize with the wavelength-dependent intensity within the incident pulse, which can be obtained either analytically from the known incidence pulse shape or from a numerical reference simulation of power measurement in an empty system.

Also in FDTD, as in FEM, PMLs can be used for representing regions extending to infinity.[154] For emission modeling, the

dipole emitter can be included directly in the system in FDTD simulations.[157,158]

5. Discrete dipole approximation (DDA)

The discrete dipole approximation (DDA) has received its name from the replacement of each scattering object by a discrete set of dipoles where the dipole moments are chosen to represent scattering from the replaced bulk volume of the scatterer.[159] The scattering from a dipole is known analytically, and the discrete set of dipoles couple to each other. This coupling gives rise to a linear equation set, with the solution of this coupled equation set yielding the solution for the overall light scattering problem. For additional technical details, including the mathematical foundation of the DDA and how it connects to Green's functions, the readers are referred, e.g., to the extensive review in Ref. 159.

6. Boundary element method (BEM)

Above, TMM and FMM are eigenmode methods in which eigenmodes are used for propagating the electromagnetic fields through z-invariant regions. FEM and FDTD on the other hand solve for the equations in the whole simulation volume, and the DDA distributes dipoles into the scatterer volume. In contrast, the boundary element method (BEM) is a surface integral approach where the Maxwell equations are formulated in a boundary integral form on the interfaces that separate two materials.[160–162] Typically, the boundary is discretized, yielding a matrix equation for the coupling between the discretized elements on the boundary. Thus, the BEM is a type of method of moments (MOM) approach where an infinite dimensional function space is approximated by a matrix equation in a finite dimensional subspace.[163]

After solving for the terms at the relevant boundaries, the electromagnetic field in the whole system can be obtained by propagating fields from the boundaries into the interior of the volume domains. This propagation can be achieved by propagating Green's functions from the boundaries into the homogenous domains. Such surface Green's functions are typically used already in constructing the matrix equations that are solved for,[160–162] and their use is connected with the Ewald–Oseen extinction theorem in optics.[164]

7. Green's function methods

Green's function (also known as the Green's dyadic) $\overleftrightarrow{G}(\mathbf{r}, \mathbf{r_s}, \lambda)$ contains extensive system information. It is particularly useful in emission problems involving an arbitrary ensemble of incoherent emitters, since it is by definition linked to the electric field by $\mathbf{E}(\mathbf{r}, \lambda) = i\omega\mu \int \overleftrightarrow{G}(\mathbf{r}, \mathbf{r_s}, \lambda)\mathbf{J}(\mathbf{r_s}, \lambda)\, d\mathbf{r_s}$ [see, e.g., Eq. (2.79) in Ref. 22]. Also, the LDOS at $\mathbf{r}_s$ and λ can be obtained from the imaginary part of $\overleftrightarrow{G}(\mathbf{r_s}, \mathbf{r_s}, \lambda)$.[22] Therefore, knowledge of $\overleftrightarrow{G}(\mathbf{r}, \mathbf{r_s}, \lambda)$ is very valuable for the design of light-emitting structures and the analysis of their underlying optical response. However, it is typically not trivial to find $\overleftrightarrow{G}(\mathbf{r}, \mathbf{r_s}, \lambda)$ for geometries beyond the homogenous medium. For a system consisting of two semi-infinite half spaces separated by a planar interface, $\overleftrightarrow{G}(\mathbf{r}, \mathbf{r_s}, \lambda)$ can be obtained analytically (see section 10.4 in Ref. 22). Developed computational schemes for the calculation of $\overleftrightarrow{G}(\mathbf{r}, \mathbf{r_s}, \lambda)$ exist, for example, with finite-difference[165] and finite-element[166] approaches. For layered media, including periodic in-plane patterning, plane waves combined with a transfer matrix scheme can be used for calculating the Green's function (see, e.g., Refs. 167 and 168).

VII. COMMENTS ON THE USE OF THE DIFFERENT METHODS

Typically, if an analytical solution exists for a given light scattering or emission problem, it is highly recommended to use it in order to minimize simulation time and to achieve better accuracy. Arguably, the most elegant means of dealing with the Maxwell equations is with Green's function approach. If possible to construct with analytical or numerical means, Green's function for a system can be used over and over again to calculate the generated fields from different source distributions, saving precious computing resources and simulation time.

As discussed partly in Secs. IV B 1 and V C, common applications for a planar optical stack include anti-reflection coatings, distributed Bragg reflectors, and optical filters. A distributed Bragg reflector stack aims to maximize reflection for a given wavelength band and is composed of alternating layers of different refractive indices. The layers in a stack-based optical filter are chosen such that either the reflection or transmission of the stack is minimized in the target wavelength band. For any of these applications, the TMM is sufficient and convenient.

FMM, FEM, and FDTD allow readily the implementation of periodic boundary conditions, which has made them the three most popular methods for modeling periodic nanostructures (it should be noted here that with periodic boundary conditions, it is enough to consider just a single unit cell of the periodic structure). FEM and FDTD yield directly information about the spatially resolved fields. FMM, on the other hand, yields direct information of the excitation of eigenmodes of the system.[21] For a recent comparison of FMM, FEM, and FDTD for modeling absorption in periodic semiconductor nanostructure arrays, the readers are encouraged to see Ref. 21. In that study, it was concluded that the preferential method depends on (i) the exact type of array solved for, even if the considered arrays resembled each other closely, and (ii) on the output sought for from the simulations (e.g., overall absorption or spatially resolved absorption).

The main drawback of FMM is its scaling with unit cell area in the in-plane directions (where the plane wave basis is used for determining the eigenmodes). Typically, the number N of plane waves needed scales with this unit cell area, and the computational time scales approximately as N^3 and the random access memory (RAM) usage as N^2.[21] However, for a small unit cell area, FMM can outperform FEM and FDTD by orders of magnitude in computational time and RAM usage.[21]

Concerning optics simulations, aperiodic structures, like metasurfaces, can typically be modeled with a supercell approach, that is, with a unit cell with a large-area in the in-plane directions. Our testing indicated that FDTD might be most suitable as compared to FMM and FEM for such structures.[21]

In FDTD, which works in the time-domain, it is not possible to use exactly the same broadband refractive index values as tabulated from experimental measurements (instead, analytical fitting

needs to be done). Therefore, discrepancies can emerge between the results obtained from FDTD and other methods, such as FMM or FEM, that work in the wavelength domain where tabulated values can be directly used.[21] On the other hand, the time-domain nature of FDTD allows us to cover a wide range of wavelengths in a single simulation run.[21]

One of the strengths of FEM is its natural ability for adaptive meshing to finely resolve physical regions where strong gradients in the electric field show up (particularly at a sharp tip of metallic particles where the lightning rod effect can show up, in the gap between two close lying metallic nanoparticles, or generally in boundary corners). However, for periodic systems, the memory requirement of FEM appears rather high compared to FDTD, and for small unit cell area also compared to FMM.[21]

In FEM and FDTD, PMLs can be used in one, two, or three dimensions. PMLs in all three dimensions allow modeling of a finite number of nanoparticles or other nanostructures. Also in FMM, it is possible to employ PMLs to model regions extending to infinity perpendicular to the propagation direction of the eigenmodes.[169] Formally, in FMM, it is thus possible to model scattering from even a single nanoparticle by choosing the direction of the incident light as the direction along which eigenmodes are solved for and using PMLs in the other two directions. However, we believe that the use of such PMLs in FMM is more efficient for analysis of waveguides, for example, in integrated optics where the waveguide modes then show up as the eigenmodes in FMM.[169]

DDA and BEM are popular for analyzing nanoparticles.[162] DDA appears to be promising for structures where only a limited scatterer volume needs to be represented by the discrete dipoles. Similarly, BEM discretizes only the surface where the refractive index changes between materials. BEM can be implemented for periodic structures, see, e.g., Ref. 170. Also, DDA supports periodic structures,[171] but in our initial testing, its numerical cost for modeling absorption in a semiconductor nanowire array was considerably higher as compared to FMM, FEM, and FDTD.

Regarding numerical implementations of the methods, for Green's function method in layered grating structures, we have used an in-house developed implementation.[167] For FMM, S^4 is a popular open source implementation,[172] but we have again relied on our in-house developed implementations.[20,173] For FDTD, we have employed the commercial software Lumerical FDTD solutions[21] and also tested the open source package MEEP.[174] For FEM, we have used exclusively the commercial software COMSOL Multiphysics, relying on its Wave Optics module, which gives built-in support for Maxwell equations, including PMLs.[21,26] For DDA, DDSCAT is a popular open source implementation,[175] which we have used for single nanowires.[176]

VIII. COMMENTS ON MORE ADVANCED MATERIAL RESPONSE

In this Tutorial, we focused on an optical response of the constituent materials that was fully described by $n(\lambda)$, with $n(\lambda)$ being the values for the corresponding bulk material. As discussed above, such type of response governs a large range of materials and applications of interest. However, by allowing for a more complicated optical response, new avenues for nanophotonics open up. First, even though we focused on the linear type of optical response (which is at the heart of many optics applications), it is worthwhile to mention that there is noticeable interest in controlling and designing the nonlinear optical response in and with nanostructures.[177–180]

Some materials show an anisotropic refractive index (e.g., III-nitride semiconductors which due to the symmetry of their wurtzite crystal structure induce directional dependence in the refractive index). Such anisotropic optical response is hence described by a 3×3 refractive index tensor $\hat{n}(\mathbf{r}, \lambda)$. Most simulation methods allow the use of such anisotropic materials in modeling with just a minor modification to the solution formulation, by including $\hat{n}(\mathbf{r}, \lambda)$ in the governing equations.

At very small length scale, the use of the (macroscopic) Maxwell equations with the bulk refractive index could be questionable. For example, in very closely lying plasmonic nanoparticles, effects from quantum mechanical confinement and tunneling of the electrons in the electron cloud can strongly modify the optical response, as compared to the predictions from the conventional macroscopic Maxwell equations.[181]

We focused on transparent materials (where $\mathrm{Im}[n(\mathbf{r}, \lambda)] = 0$) and absorbing materials (where $\mathrm{Im}[n(\mathbf{r}, \lambda)] > 0$). However, in various application, e.g., in optical amplifiers[182] and lasers,[183] active materials with $\mathrm{Im}[n(\mathbf{r}, \lambda)] < 0$, giving rise to optical gain, are used. In absorption modeling, we typically assume that the absorption is weak enough to not saturate the capability of the material to absorb further, hence assuming a constant $\mathrm{Im}[n(\mathbf{r}, \lambda)] > 0$. However, in amplifiers and lasers, the amplified optical field could start to deplete the gain available from the underlying active medium. In such a case, the assumption of a constant $\mathrm{Im}[n(\mathbf{r}, \lambda)] < 0$ is not necessarily appropriate. One approach to simplify the analysis is to set $\mathrm{Im}[n(\mathbf{r}, \lambda)] = 0$ for the gain medium, after which the optical modes of the structure are solved for, yielding valuable information for the design of high-gain nanostructures.[183]

The Lorentz reciprocity, which relates the response from one emitter to another (as discussed in Sec. VI B 2), holds for materials where ε and μ are symmetric tensors (including thus symmetric scalars) that are time-independent.[184] Thus, by inducing either non-symmetry or time-dependence to ε or μ, we can break the Lorentz reciprocity. With such Lorentz reciprocity breaking materials, it would be possible to construct optical isolators that allow transmission of light only in one direction, which would thus function as the equivalent of a diode for integrated optics.[184]

In this Tutorial, we focused on the fully coherent Maxwell equations. Thus, in that model, we find interference for light reflected from two interfaces, no matter the distance between the interfaces. Such an assumption is typically appropriate if we focus on diffraction from small-sized nanostructures or from a thin nanostructured layer (or on light from a light source with a long coherence length, such as a continuous wave laser). However, in some systems, the coherence of light can be lost between more remote interfaces, particularly if considering thermal radiation (such as sunlight) that has a limited coherence length. For more information on how to study such systems, see, for example, Ref. 185 that presents the modeling of a Si solar cell with a nanostructured front and rear interface. There, the fully coherent Maxwell equations are solved for at each of the interfaces, but light is propagated in an incoherent fashion through the thick Si absorber.

Here, we focused on bulk-like materials. In contrast, the emerging 2D materials are essentially crystalline solids with atoms arranged in a 2D lattice. The reduced dimensionality open up new avenues for photonics.[178,186–188] The most well-known and studied 2D material is graphene, which is composed of carbon atoms in a 2D hexagonal lattice (the multi-layer bulk form of graphene is known as graphite). Other notable 2D materials include several transition metal dichalcogenides, hexagonal boron nitride, molybdenum disulfide, and black phosphorus. The planar structure of 2D materials and resulting quantum confinement effects in the direction perpendicular to the lattice plane lead to many extraordinary electronic and optical properties. A database is available, listing optical and several other properties for a large number of 2D materials.[189] For example, graphene has a high optical transparency, yet relatively large absorption coefficient, and a mostly flat, wide absorption spectrum reaching from around 300 to 2500 nm in wavelength. Many other 2D materials exhibit a bandgap making them insulators or semiconductors while certain surface treatments can open a bandgap in graphene.[186]

Finally, throughout this Tutorial, we have taken a classical footing on optics. However, particularly when moving to the level of single or few photons in the optical field, additional quantum mechanical aspects of light can become important.[190] In such systems, effects like squeezed light,[191] entanglement,[192] and quantum coherence[193] start to appear, opening up a manifold of possible photonics applications in quantum information processing.[194]

IX. CONCLUDING REMARKS

The aim of this Tutorial is to provide the reader with inspiring information on the many possibilities that nanostructures bring to photonics and on the avenues opened by computational electromagnetic optics for the design and analysis of nanophotonics applications. We recommend the testing and continuous use of multiple simulation methods to gain hands-on experience of their strengths and weaknesses for varying types of materials, geometries, and intended applications.

ACKNOWLEDGMENTS

This work is supported by the Academy of Finland [Grant No. 320166 (PREIN Flagship—University of Eastern Finland) and Grant No. 320167 (PREIN Flagship—Aalto University)]. We acknowledge the computational resources provided by the Aalto Science-IT project.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]G. G. Kang, I. Vartiainen, B. F. Bai, H. Tuovinen, and J. Turunen, Appl. Phys. Lett. **99**, 071103 (2011).
[2]T. Pasanen, V. Vähänissi, N. Theut, and H. Savin, Energy Procedia **124**, 307 (2017).
[3]S. Guldin, S. Hüttner, M. Kolle, M. E. Welland, P. Müller-Buschbaum, R. H. Friend, U. Steiner, and N. Tétreault, Nano Lett. **10**, 2303 (2010).
[4]J. Wallentin, N. Anttu, D. Asoli, M. Huffman, I. Åberg, M. H. Magnusson, G. Siefer, P. Fuss-Kailuweit, F. Dimroth, B. Witzigmann, H. Q. Xu, L. Samuelson, K. Deppert, and M. T. Borgström, Science **339**, 1057 (2013).
[5]S. A. Khan, D. B. Suyatin, J. Sundqvist, M. Graczyk, M. Junige, C. Kauppinen, A. Kvennefors, M. Huffman, and I. Maximov, ACS Appl. Nano Mater. **1**, 2476 (2018).
[6]H. Husu, B. K. Canfield, J. Laukkanen, B. Bai, M. Kuittinen, J. Turunen, and M. Kauranen, Appl. Phys. Lett. **93**, 183115 (2008).
[7]E. Barrigón, M. Heurlin, Z. Bi, B. Monemar, and L. Samuelson, Chem. Rev. **119**, 9170 (2019).
[8]B. Ai, H. Möhwald, D. Wang, and G. Zhang, Adv. Mater. Interfaces **4**, 1600271 (2017).
[9]L. Romano and M. Stampanoni, Micromachines **11**, 589 (2020).
[10]Y. Zhao, X. Dai, F. Wang, X. Zhang, C. Fan, and X. Liu, Nano Today **26**, 123 (2019).
[11]N. Chekurov, K. Grigoras, A. Peltonen, S. Franssila, and I. Tittonen, Nanotechnology **20**, 065307 (2009).
[12]Y. Chen, Microelectron. Eng. **135**, 57 (2015).
[13]H. Schift, J. Vac. Sci. Technol. B **26**, 458 (2008).
[14]P.-I. Schneider, X. Garcia Santiago, V. Soltwisch, M. Hammerschmidt, S. Burger, and C. Rockstuhl, ACS Photonics **6**, 2726 (2019).
[15]L. Chayanun, L. Hrachowina, A. Björling, M. T. Borgström, and J. Wallentin, Nano Lett. **20**, 8326 (2020).
[16]V. Giannini, A. Berrier, S. A. Maier, J. A. Sánchez-Gil, and J. G. Rivas, Opt. Express **18**, 2797 (2010).
[17]B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, 2nd ed. (John Wiley & Sons, 2012).
[18]J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (John Wiley & Sons, 1999).
[19]N. Anttu, M. Heurlin, M. T. Borgström, M.-E. Pistol, H. Q. Xu, and L. Samuelson, Nano Lett. **13**, 2662 (2013).
[20]N. Anttu and H. Q. Xu, Phys. Rev. B **83**, 165431 (2011).
[21]N. Anttu, H. Mäntynen, T. Sadi, A. Matikainen, J. Turunen, and H. Lipsanen, Nano Express **1**, 030034 (2020).
[22]L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, 2006).
[23]G. Grzela, R. Paniagua-Domínguez, T. Barten, Y. Fontana, J. A. Sánchez-Gil, and J. Gómez Rivas, Nano Lett. **12**, 5481 (2012).
[24]K. Joulain, R. Carminati, J.-P. Mulet, and J.-J. Greffet, Phys. Rev. B **68**, 245405 (2003).
[25]M. Pelton, Nat. Photonics **9**, 427 (2015).
[26]N. Anttu, H. Mäntynen, A. Sorokina, P. Kivisaari, T. Sadi, and H. Lipsanen, Photonics **7**, 23 (2020).
[27]O. L. Muskens, V. Giannini, J. A. Sánchez-Gil, and J. Gómez Rivas, Nano Lett. **7**, 2871 (2007).
[28]N. Anttu, Opt. Lett. **41**, 1494 (2016).
[29]P. Kivisaari, Y. Chen, and N. Anttu, Nano Futures **2**, 015001 (2018).
[30]R. Buschlinger, M. Lorke, and U. Peschel, Phys. Rev. B **91**, 045203 (2015).
[31]E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, 1985), pp. 3–9.
[32]I. H. Malitson, J. Opt. Soc. Am. **55**, 1205 (1965).
[33]J. R. DeVore, J. Opt. Soc. Am. **41**, 416 (1951).
[34]P. B. Johnson and R. W. Christy, Phys. Rev. B **6**, 4370 (1972).
[35]M. A. Green, Sol. Energy Mater. Sol. Cells **92**, 1305 (2008).
[36]E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, 1985), pp. 429–444.
[37]I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001).
[38]C.-Y. Tsai, J. Appl. Phys. **123**, 183103 (2018).
[39]N. Anttu, Opt. Lett. **38**, 730 (2013).
[40]S. Daqiqeh Rezaei, Z. Dong, J. You En Chan, J. Trisno, R. J. H. Ng, Q. Ruan, C.-W. Qiu, N. A. Mortensen, and J. K. W. Yang, ACS Photonics **8**, 18 (2021).
[41]C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (John Wiley & Sons, 2008).

[42]V. Amendola, R. Pilot, M. Frasconi, O. M. Maragò, and M. A. Iatì, J. Phys.: Condens. Matter **29**, 203002 (2017).
[43]A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, and B. Luk'yanchuk, Science **354**, aag2472 (2016).
[44]G. Grinblat, Y. Li, M. P. Nielsen, R. F. Oulton, and S. A. Maier, Nano Lett. **16**, 4635 (2016).
[45]E. Prodan, C. Radloff, N. J. Halas, and P. Nordlander, Science **302**, 419 (2003).
[46]M. Hentschel, M. Saliba, R. Vogelgesang, H. Giessen, A. P. Alivisatos, and N. Liu, Nano Lett. **10**, 2721 (2010).
[47]B. Gerislioglu, L. Dong, A. Ahmadivand, H. Hu, P. Nordlander, and N. J. Halas, Nano Lett. **20**, 2087 (2020).
[48]A. E. Miroshnichenko and Y. S. Kivshar, Nano Lett. **12**, 6459 (2012).
[49]M. F. Limonov, M. V. Rybin, A. N. Poddubny, and Y. S. Kivshar, Nat. Photonics **11**, 543 (2017).
[50]J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Mead, *Photonic Crystals: Molding the Flow of Light* (Princeton University Press, 2008).
[51]E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987).
[52]T. Baba, Nat. Photonics **2**, 465 (2008).
[53]H. Savin, P. Repo, G. von Gastrow, P. Ortega, E. Calle, M. Garín, and R. Alcubilla, Nat. Nanotechnol. **10**, 624 (2015).
[54]H. Fredriksson, Y. Alaverdyan, A. Dmitriev, C. Langhammer, D. S. Sutherland, M. Zäch, and B. Kasemo, Adv. Mater. **19**, 4297 (2007).
[55]D. S. Wiersma, P. Bartolini, A. Lagendijk, and R. Righini, Nature **390**, 671 (1997).
[56]H. Cao, Waves Random Media **13**, R1 (2003).
[57]J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company Publishers, 2005).
[58]P. Qiao, L. Zhu, W. C. Chew, and C. J. Chang-Hasnain, Opt. Express **23**, 24508 (2015).
[59]M. Neviere and E. K. Popov, *Theory and Practice of Surface-Relief Diffraction Gratings: Synchrotron and Other Applications* (International Society for Optics and Photonics, 1998), pp. 2–10.
[60]M. Neviere and E. Popov, *Light Propagation in Periodic Media: Differential Theory and Design* (CRC Press, 2002).
[61]H. Kim, J. Park, and B. Lee, *Fourier Modal Method and Its Applications in Computational Nanophotonics* (CRC Press, 2017).
[62]S. Fan and J. D. Joannopoulos, Phys. Rev. B **65**, 235112 (2002).
[63]P. Lalanne, J. P. Hugonin, and P. Chavel, J. Lightwave Technol. **24**, 2442 (2006).
[64]V. Dagytė and N. Anttu, Nanotechnology **30**, 025710 (2018).
[65]H. P. Herzig, *Micro-Optics: Elements, Systems and Applications* (CRC Press, 1997).
[66]J. Turunen and F. Wyrovski, *Diffractive Optics* (Akademie, Berlin, 1997).
[67]J. Turunen, M. Kuittinen, and F. Wyrowski, in *Progress in Optics*, edited by E. Wolf (Elsevier, 2000), pp. 343–388.
[68]*Metasurfaces: Physics and Applications*, edited by S. I. Bozhevolnyi, P. Genevet, and F. Ding (MDPI, 2018).
[69]P. B. Clapham and M. C. Hutley, Nature **244**, 281 (1973).
[70]*Electromagnetic Theory of Gratings*, edited by R. Petit (Springer-Verlag, 1980).
[71]T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, Nature **391**, 667 (1998).
[72]M. W. Farn, Appl. Opt. **31**, 4453 (1992).
[73]J. M. Finlan, K. M. Flood, and R. J. Bojko, Opt. Eng. **34**, 3560 (1995).
[74]S. Banerji, M. Meem, A. Majumder, F. G. Vasquez, B. Sensale-Rodriguez, and R. Menon, Optica **6**, 805 (2019).
[75]N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, Science **334**, 333 (2011).
[76]N. Yu and F. Capasso, Nat. Mater. **13**, 139 (2014).
[77]M. Khorasaninejad, W. T. Chen, R. C. Devlin, J. Oh, A. Y. Zhu, and F. Capasso, Science **352**, 1190 (2016).
[78]O. Quevedo-Teruel, H. Chen, A. Díaz-Rubio, G. Gok, A. Grbic, G. Minatti, E. Martini, S. Maci, G. V. Eleftheriades, M. Chen, N. I. Zheludev, N. Papasimakis, S. Choudhury, Z. A. Kudyshev, S. Saha, H. Reddy, A. Boltasseva, V. M. Shalaev, A. V. Kildishev, D. Sievenpiper, C. Caloz, A. Alù, Q. He, L. Zhou, G. Valerio, E. Rajo-Iglesias, Z. Sipus, F. Mesa, R. Rodríguez-Berral, F. Medina, V. Asadchy, S. Tretyakov, and C. Craeye, J. Opt. **21**, 073002 (2019).
[79]A. Poddubny, I. Iorsh, P. Belov, and Y. Kivshar, Nat. Photonics **7**, 948 (2013).
[80]D. R. Smith, J. B. Pendry, and M. C. K. Wiltshire, Science **305**, 788 (2004).
[81]J. B. Pendry, D. Schurig, and D. R. Smith, Science **312**, 1780 (2006).
[82]N. I. Zheludev and Y. S. Kivshar, Nat. Mater. **11**, 917 (2012).
[83]S. Kruk and Y. Kivshar, ACS Photonics **4**, 2638 (2017).
[84]Y. H. Ko and R. Magnusson, Optica **5**, 289 (2018).
[85]P. Wang, M. E. Nasir, A. V. Krasavin, W. Dickson, Y. Jiang, and A. V. Zayats, Acc. Chem. Res. **52**, 3018 (2019).
[86]J. Cai and L. Qi, Mater. Horiz. **2**, 37 (2015).
[87]L. Novotny and N. van Hulst, Nat. Photonics **5**, 83 (2011).
[88]I. Friedler, C. Sauvan, J. P. Hugonin, P. Lalanne, J. Claudon, and J. M. Gérard, Opt. Express **17**, 2095 (2009).
[89]G. Bulgarini, M. E. Reimer, M. Bouwes Bavinck, K. D. Jöns, D. Dalacu, P. J. Poole, E. P. A. M. Bakkers, and V. Zwiller, Nano Lett. **14**, 4102 (2014).
[90]B. Rolly, B. Stout, and N. Bonod, Opt. Express **20**, 20376 (2012).
[91]M. S. Eggleston, K. Messer, L. Zhang, E. Yablonovitch, and M. C. Wu, Proc. Natl. Acad. Sci. **112**, 1704 (2015).
[92]L. Tang, S. E. Kocabas, S. Latif, A. K. Okyay, D.-S. Ly-Gagnon, K. C. Saraswat, and D. A. B. Miller, Nat. Photonics **2**, 226 (2008).
[93]R.-C. Ge and S. Hughes, Opt. Lett. **39**, 4235 (2014).
[94]K. J. Vahala, Nature **424**, 839 (2003).
[95]Q. Li, T. Wang, Y. Su, M. Yan, and M. Qiu, Opt. Express **18**, 8367 (2010).
[96]M. R. Foreman, J. D. Swaim, and F. Vollmer, Adv. Opt. Photonics **7**, 168 (2015).
[97]I. Aharonovich, D. Englund, and M. Toth, Nat. Photonics **10**, 631 (2016).
[98]H. Mäntynen, N. Anttu, Z. Sun, and H. Lipsanen, Nanophotonics **8**, 747 (2019).
[99]P.-I. Schneider, N. Srocka, S. Rodt, L. Zschiedrich, S. Reitzenstein, and S. Burger, Opt. Express **26**, 8479 (2018).
[100]N. Gregersen, T. R. Nielsen, J. Mørk, J. Claudon, and J.-M. Gérard, Opt. Express **18**, 21204 (2010).
[101]I. Schnitzer, E. Yablonovitch, C. Caneau, T. J. Gmitter, and A. Scherer, Appl. Phys. Lett. **63**, 2174 (1993).
[102]T. Sadi, I. Radevici, and J. Oksanen, Nat. Photonics **14**, 205 (2020).
[103]P. T. Landsberg, *Recombination in Semiconductors* (Cambridge University Press, 2003).
[104]J. Piprek, Materials **13**, 5174 (2020).
[105]A. I. Zhmakin, Phys. Rep. **498**, 189 (2011).
[106]P. Kivisaari, L. Riuttanen, J. Oksanen, S. Suihkonen, M. Ali, H. Lipsanen, and J. Tulkki, Appl. Phys. Lett. **101**, 021113 (2012).
[107]J. Piprek, Phys. Status Solidi RRL **8**, 424 (2014).
[108]C. Wiesmann, K. Bergenek, N. Linder, and U. T. Schwarz, Laser Photonics Rev. **3**, 262 (2009).
[109]X. Jin, S. Trieu, G. J. Chavoor, and G. M. Halpin, Nanomaterials **8**, 1045 (2018).
[110]G. Lozano, S. R. Rodriguez, M. A. Verschuuren, and J. G. Rivas, Light: Sci. Appl. **5**, e16080 (2016).
[111]T. Sadi, J. Oksanen, and J. Tulkki, J. Appl. Phys. **114**, 223104 (2013).
[112]REN21, Renewable Global Status Report Technical Report, REN21 Secretariat, Paris, France, 2020.
[113]J. F. Geisz, R. M. France, K. L. Schulte, M. A. Steiner, A. G. Norman, H. L. Guthrey, M. R. Young, T. Song, and T. Moriarty, Nat. Energy **5**, 326 (2020).
[114]M. Green, E. Dunlop, J. Hohl-Ebinger, M. Yoshita, N. Kopidakis, and X. Hao, "Solar cell efficiency tables (version 57)," Prog. Photovoltaics **29**, 3–15 (2021)
[115]Z. Yu, A. Raman, and S. Fan, Proc. Natl. Acad. Sci. U.S.A. **107**, 17491 (2010).
[116]G. Otnes and M. T. Borgström, Nano Today **12**, 31 (2017).
[117]E. Yablonovitch, J. Opt. Soc. Am. **72**, 899 (1982).
[118]S. Mokkapati and K. R. Catchpole, J. Appl. Phys. **112**, 101101 (2012).
[119]C. Battaglia, C.-M. Hsu, K. Söderström, J. Escarré, F.-J. Haug, M. Charrière, M. Boccard, M. Despeisse, D. T. L. Alexander, M. Cantoni, Y. Cui, and C. Ballif, ACS Nano **6**, 2790 (2012).

[120]L. Cao, J. S. White, J.-S. Park, J. A. Schuller, B. M. Clemens, and M. L. Brongersma, Nat. Mater. **8**, 643 (2009).
[121]P. Krogstrup, H. I. Jørgensen, M. Heiss, O. Demichel, J. V. Holm, M. Aagesen, J. Nygard, and A. Fontcuberta i Morral, Nat. Photonics **7**, 306 (2013).
[122]I.-J. Chen, S. Limpert, W. Metaferia, C. Thelander, L. Samuelson, F. Capasso, A. M. Burke, and H. Linke, Nano Lett. **20**, 4064 (2020).
[123]E. Garnett and P. Yang, Nano Lett. **10**, 1082 (2010).
[124]C. Lin and M. L. Povinelli, Opt. Express **17**, 19371 (2009).
[125]J. A. Czaban, D. A. Thompson, and R. R. LaPierre, Nano Lett. **9**, 148 (2009).
[126]J. Kupec and B. Witzigmann, Opt. Express **17**, 10399 (2009).
[127]K. Seo, M. Wober, P. Steinvurzel, E. Schonbrun, Y. Dan, T. Ellenbogen, and K. B. Crozier, Nano Lett. **11**, 1851 (2011).
[128]Y. Cui, J. Wang, S. R. Plissard, A. Cavalli, T. T. T. Vu, R. P. J. van Veldhoven, L. Gao, M. Trainor, M. A. Verheijen, J. E. M. Haverkort, and E. P. A. M. Bakkers, Nano Lett. **13**, 4113 (2013).
[129]D. van Dam, N. J. J. van Hoof, Y. Cui, P. J. van Veldhoven, E. P. A. M. Bakkers, J. Gómez Rivas, and J. E. M. Haverkort, ACS Nano **10**, 11414 (2016).
[130]S. L. Diedenhofen, O. T. A. Janssen, G. Grzela, E. P. A. M. Bakkers, and J. Gómez Rivas, ACS Nano **5**, 2316 (2011).
[131]I. Åberg, G. Vescovi, D. Asoli, U. Naseem, J. P. Gilboy, C. Sundvall, A. Dahlgren, K. E. Svensson, N. Anttu, M. T. Björk, and L. Samuelson, IEEE J. Photovolt. **6**, 185 (2016).
[132]G. Otnes, E. Barrigón, C. Sundvall, K. E. Svensson, M. Heurlin, G. Siefer, L. Samuelson, I. Åberg, and M. T. Borgström, Nano Lett. **18**, 3038 (2018).
[133]S. J. Gibson, B. van Kasteren, B. Tekcan, Y. Cui, D. van Dam, J. E. M. Haverkort, E. P. A. M. Bakkers, and M. E. Reimer, Nat. Nanotechnol. **14**, 473 (2019).
[134]V. Raj, K. Vora, L. Fu, H. H. Tan, and C. Jagadish, ACS Nano **13**, 12015 (2019).
[135]N. Anttu and H. Q. Xu, Opt. Express **21**, A558 (2013).
[136]I. Radevici, J. Tiira, T. Sadi, S. Ranta, A. Tukiainen, M. Guina, and J. Oksanen, Appl. Phys. Lett. **114**, 051101 (2019).
[137]B. Zhao, P. Santhanam, K. Chen, S. Buddhiraju, and S. Fan, Nano Lett. **18**, 5224 (2018).
[138]A. I. Mahan, J. Opt. Soc. Am. **46**, 913 (1956).
[139]C. C. Katsidis and D. I. Siapkas, Appl. Opt. **41**, 3978 (2002).
[140]J. V. Dave, IBM J. Res. Dev. **13**, 302 (1969).
[141]L. Rayleigh, Philos. Mag. **47**, 375 (1899).
[142]M. I. Mishchenko, L. D. Travis, and D. W. Mackowski, J. Quant. Spectrosc. Radiat. Transfer **55**, 535 (1996).
[143]V. A. Markel, J. Opt. Soc. Am. A **33**, 1244 (2016).
[144]V. S. Asadchy, M. S. Mirmoosa, A. Díaz-Rubio, S. Fan, and S. A. Tretyakov, Proc. IEEE **108**, 1684 (2020).
[145]J. Yang, J.-P. Hugonin, and P. Lalanne, ACS Photonics **3**, 395 (2016).
[146]E. Hecht, *Optics* (Addison-Wesley, 2002).
[147]S. F. Helfert and R. Pregla, Electromagnetics **22**, 615 (2002).
[148]D. Y. K. Ko and J. C. Inkson, Phys. Rev. B **38**, 9945 (1988).
[149]D. M. Whittaker and I. S. Culshaw, Phys. Rev. B **60**, 2610 (1999).
[150]L. Li, J. Opt. Soc. Am. A **13**, 1024 (1996).
[151]E. Popov, M. Nevière, B. Gralak, and G. Tayeb, J. Opt. Soc. Am. A **19**, 33 (2002).
[152]P. Monk, *Finite Element Methods for Maxwell's Equations* (Clarendon Press, 2003).
[153]C. M. Rappaport and B. J. McCartin, IEEE Trans. Antennas Propag. **39**, 345 (1991).
[154]J.-P. Berenger, J. Comput. Phys. **114**, 185 (1994).
[155]K. Yee, IEEE Trans. Antennas Propag. **14**, 302 (1966).
[156]A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Artech House, 2005).
[157]A. Didari and M. P. Mengüç, J. Quant. Spectrosc. Radiat. Transfer **146**, 214 (2014).
[158]M. H. Chowdhury, J. Pond, S. K. Gray, and J. R. Lakowicz, J. Phys. Chem. C **112**, 11236 (2008).
[159]M. A. Yurkin and A. G. Hoekstra, J. Quant. Spectrosc. Radiat. Transfer **106**, 558 (2007).
[160]F. J. García de Abajo and A. Howie, Phys. Rev. B **65**, 115418 (2002).
[161]P. Yla-Oijala, J. Markkanen, S. Jarvenpaa, and S. P. Kiminki, Prog. Electromagn. Res. **149**, 15 (2014).
[162]V. Myroshnychenko, J. Rodríguez-Fernández, I. Pastoriza-Santos, A. M. Funston, C. Novo, P. Mulvaney, L. M. Liz-Marzán, and F. J. G. de Abajo, Chem. Soc. Rev. **37**, 1792 (2008).
[163]R. F. Harrington, J. Electromagn. Waves Appl. **1**, 181 (1987).
[164]A. Lakhtakia, in *The World of Applied Electromagnetics: In Appreciation of Magdy Fahmy Iskander*, edited by A. Lakhtakia and C. M. Furse (Springer International Publishing, Cham, 2018), pp. 481–513.
[165]A. J. Ward and J. B. Pendry, Phys. Rev. B **58**, 7252 (1998).
[166]W. Li, D. Tan, J. Xu, S. Wang, and Y. Chen, Opt. Express **27**, 16047 (2019).
[167]T. Sadi, J. Oksanen, J. Tulkki, P. Mattila, and J. Bellessa, IEEE J. Sel. Top. Quantum Electron. **19**, 1 (2013).
[168]J. E. Sipe, J. Opt. Soc. Am. B **4**, 481 (1987).
[169]E. Silberstein, P. Lalanne, J.-P. Hugonin, and Q. Cao, J. Opt. Soc. Am. A **18**, 2865 (2001).
[170]B. Gallinet, A. M. Kern, and O. J. F. Martin, J. Opt. Soc. Am. A **27**, 2261 (2010).
[171]B. T. Draine and P. J. Flatau, J. Opt. Soc. Am. A **25**, 2693 (2008).
[172]V. Liu and S. Fan, Comput. Phys. Commun. **183**, 2233 (2012).
[173]E. Noponen and J. Turunen, J. Opt. Soc. Am. A **11**, 2494 (1994).
[174]A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. D. Joannopoulos, and S. G. Johnson, Comput. Phys. Commun. **181**, 687 (2010).
[175]B. T. Draine and P. J. Flatau, J. Opt. Soc. Am. A **11**, 1491 (1994).
[176]Y. Chen, N. Anttu, S. Sivakumar, E. Gompou, and M. H. Magnusson, Nanotechnology **31**, 134001 (2020).
[177]D. Smirnova and Y. S. Kivshar, Optica **3**, 1241 (2016).
[178]A. Autere, H. Jussila, Y. Dai, Y. Wang, H. Lipsanen, and Z. Sun, Adv. Mater. **30**, 1705963 (2018).
[179]M. Kauranen and A. V. Zayats, Nat. Photonics **6**, 737 (2012).
[180]M. SoljaČiĆ and J. D. Joannopoulos, Nat. Mater. **3**, 211 (2004).
[181]M. S. Tame, K. R. McEnery, Ş. K. Özdemir, J. Lee, S. A. Maier, and M. S. Kim, Nat. Phys. **9**, 329 (2013).
[182]J. Rönn, W. Zhang, A. Autere, X. Leroux, L. Pakarinen, C. Alonso-Ramos, A. Säynätjoki, H. Lipsanen, L. Vivien, E. Cassan, and Z. Sun, Nat. Commun. **10**, 432 (2019).
[183]H. Mäntynen, N. Anttu, and H. Lipsanen, Materials **13**, 5510 (2020).
[184]D. Jalas, A. Petrov, M. Eich, W. Freude, S. Fan, Z. Yu, R. Baets, M. Popović, A. Melloni, J. D. Joannopoulos, M. Vanwolleghem, C. R. Doerr, and H. Renner, Nat. Photonics **7**, 579 (2013).
[185]N. Tucher, J. Eisenlohr, P. Kiefel, O. Höhn, H. Hauser, M. Peters, C. Müller, J. C. Goldschmidt, and B. Bläsi, Opt. Express **23**, A1720 (2015).
[186]F. Xia, H. Wang, D. Xiao, M. Dubey, and A. Ramasubramaniam, Nat. Photonics **8**, 899 (2014).
[187]F. Bonaccorso, Z. Sun, T. Hasan, and A. C. Ferrari, Nat. Photonics **4**, 611 (2010).
[188]F. H. L. Koppens, T. Mueller, P. Avouris, A. C. Ferrari, M. S. Vitiello, and M. Polini, Nat. Nanotechnol. **9**, 780 (2014).
[189]S. Haastrup, M. Strange, M. Pandey, T. Deilmann, P. S. Schmidt, N. F. Hinsche, M. N. Gjerding, D. Torelli, P. M. Larsen, A. C. Riis-Jensen, J. Gath, K. W. Jacobsen, J. J. Mortensen, T. Olsen, and K. S. Thygesen, 2D Mater. **5**, 042002 (2018).
[190]A. Pathak and A. Ghatak, J. Electromagn. Waves Appl. **32**, 229 (2018).
[191]R. Loudon and P. L. Knight, J. Mod. Opt. **34**, 709 (1987).
[192]R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[193]A. Streltsov, G. Adesso, and M. B. Plenio, Rev. Mod. Phys. **89**, 041003 (2017).
[194]F. Flamini, N. Spagnolo, and F. Sciarrino, Rep. Prog. Phys. **82**, 016001 (2018).