


Efficient perturbative framework for coupling of radiative and guided modes in nearly periodic surfaces

Sophie Fisher ^{*}

Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Raphaël Pestourie and Steven G. Johnson 

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



(Received 7 March 2022; accepted 21 June 2022; published 13 July 2022)

We present a semianalytical framework for computing the coupling of radiative and guided waves in slowly varying (nearly uniform or nearly periodic) surfaces, which is especially relevant to the exploitation of nonlocal effects in large-area metasurfaces. Our framework bridges a gap in the theory of slowly varying surfaces: aside from brute-force numerical simulations, current approximate methods can model *either* guided *or* radiative waves but cannot easily model their *coupling*. We solve this problem by combining two methods: the locally periodic approximation, which approximates radiative scattering by composing a set of periodic scattering problems, and spatial coupled-wave theory, which allows the perturbative modeling of guided waves using an eigenmode expansion. We derive our framework for both nearly uniform and nearly periodic surfaces, and we validate each case against brute-force finite-difference time-domain simulations, which show increasing agreement as the surface varies more slowly.

DOI: [10.1103/PhysRevA.106.013507](https://doi.org/10.1103/PhysRevA.106.013507)

I. INTRODUCTION

Although many perturbative techniques are available to study *guided* modes propagating *through* slowly varying (nearly uniform or nearly periodic) media [1–4] and, conversely, a number of approximations have been devised for *radiative* waves scattering *off* of slowly varying surfaces [5–21], the *coupling* of guided and radiative modes by slowly varying structures is relatively unstudied except by brute-force numerical simulations [22–24]. This problem is especially relevant to large-area optical metasurfaces (comprising thin aperiodic subwavelength-scale patterns), which have been widely used for free-space wave-front engineering [25]: the large diameter of metasurface devices (often >1000 wavelengths λ) can make them impractical to design without mathematical approximations [26–34]. Such surfaces are often modeled using a locally periodic approximation (LPA), in which the far-field scattering of an incident wave from each unit cell is computed assuming a periodic surface, which works well when the unit cells are slowly varying [5–17,21]. A simplified limiting case of LPA is the locally *uniform* approximation (LUA), which corresponds to LPA in the limit as the period goes to zero: LUA approximates the structure at each point by a flat surface with a matching cross section and is also found in the form of scalar-diffraction theory [19] or the Kirchhoff tangent-plane approximation (for curved surfaces) [20]. However, LPA and LUA cannot be applied to the “nonlocal” problem of radiation coupling to and from guided modes since in a periodic surface guided modes do

not radiate by definition [35]. This is in contrast to previous nonlocal metasurfaces employing *leaky* resonances [36–43] which radiate even for purely periodic surfaces [44]. On the other hand, guided-wave propagation can be treated by spatial coupled-wave theories (CWTs)—eigenmode-expansion methods resulting in a set of coupled ordinary differential equations in the mode coefficients, which can be solved perturbatively for slowly varying media [1–4]. However, radiative modes are difficult to treat with CWT techniques because there is a continuous spectrum of radiative solutions, in contrast to a discrete, easily truncatable guided-mode basis. As a consequence, large-area metasurface designs have thus far not exploited the potential for nonlocal effects of guided modes, e.g., for long-lifetime light trapping to enhance light-matter interactions, metasurface lasers, or large-area grating couplers.

In this paper, we bridge the gap between radiation and guided modes in slowly varying (both locally periodic and locally uniform) media by a semianalytical framework combining *both* LPA or LUA and CWT, using LPA or LUA to model the incident or scattered radiation and using perturbatively coupled CWT to model the guided waves. We begin by deriving and validating our framework (against brute-force numerics) for the easier case of locally uniform metasurfaces [Fig. 1(a) and Sec. II], where the cross section of the surface varies slowly with z . We then generalize and validate our framework for the more complicated case of locally periodic metasurfaces [Fig. 1(b) and Sec. III] where the unit cell varies slowly with z , first for the case of unit cells with a fixed period and later for a z -dependent period (see the Appendix). Unlike brute-force Maxwell solvers, which require enormous resources for large-area metasurface modeling and

*sefisher@mit.edu

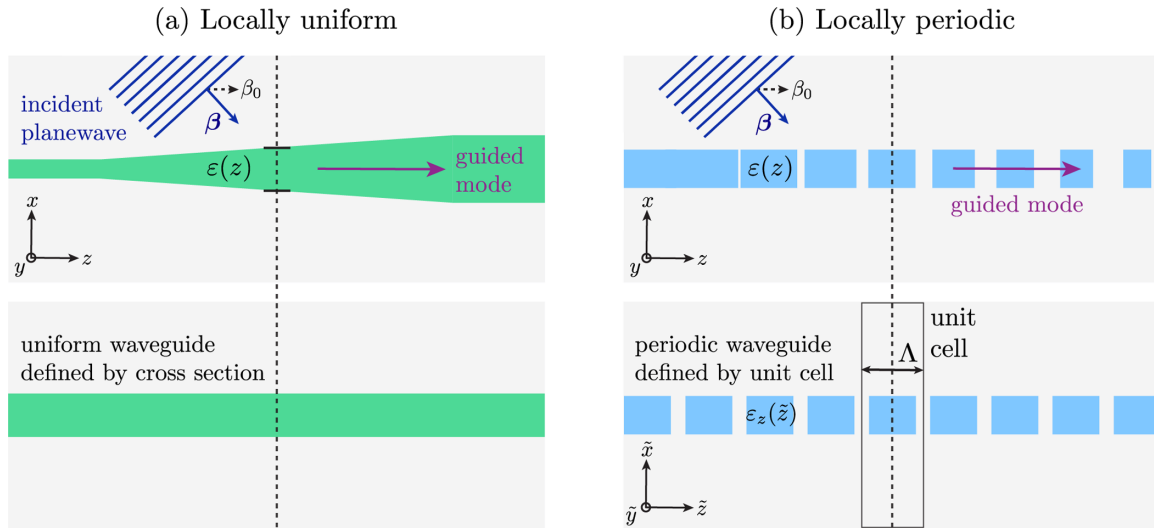


FIG. 1. (a) An incident plane wave with propagation constant β_0 impinges on a surface with locally uniform permittivity defined by $\varepsilon(z)$. The plane wave couples to guided waves propagating along the z direction. The guided waves are computed using the locally uniform approximation and coupled-wave theory (CWT), techniques that employ a basis of completely uniform waveguides defined by the cross sections of the locally uniform surface. (b) The same scattering problem as in (a), but the surface has locally *periodic* permittivity. The guided waves are computed using the locally periodic approximation and a generalized CWT for nearly periodic media. These employ a basis of *periodic* waveguides for each cross section defined by $\varepsilon_z(\tilde{z})$, with periods $\Lambda(z)$. A new coordinate, \tilde{z} , is introduced to describe the unit cells of the periodic waveguides.

optimization, our framework is computationally cheap: to calculate the coupling between an incident wave and guided mode, one computes the LPA solution and the guided-mode fields for each unit cell (periodic Maxwell solves), and then one simply evaluates an inexpensive integral of the slowly varying overlap between the LPA solution and guided-mode fields (Algorithm 1). Compared to brute-force finite-difference time-domain (FDTD) simulations, our framework shows increasing accuracy for more slowly varying structures (where FDTD becomes much more expensive) and also exhibits high accuracy even for variation on a few-wavelength scale (Figs. 2 and 3).

While brute-force numerical simulations capture effects due to both guided and radiation modes, optical metasurfaces are challenging to model in this way due to their irregular subwavelength-scale features and large diameters [25–34]. This is particularly the case for metasurface design by large-scale optimization (i.e., inverse design), which requires surfaces to be simulated many times as the geometric parameters evolve [14,17,45–52]. Ways to extend the power of brute-force Maxwell solvers to the scale of metasurfaces have been proposed, e.g., using GPU-accelerated FDTD [53,54] or spatially truncated integral-equation methods (for nontouching meta-atoms) [55]. However, these methods still face severe challenges in reaching the 10^4 -wavelength scale required for inverse design of large metasurfaces. Moreover, even when computational power reaches this scale, there will always be a place for rapid-prototyping approximations that take only a few minutes on a laptop. In fact, our framework requires the periodic Maxwell solves to be performed only *once* for a sequence of unit cells, and then they can be reused (interpolated) to optimize or explore many different slowly varying metasurfaces [4,56–60].

Many approximate techniques to model metasurfaces have already been developed. One such technique is LPA, which composes a set of periodic scattering problems for each unit cell and has been used for optimization-based inverse design [14,17,50–52] as well as selecting the phase profile of the surface *a priori* [5–11,13,16,21]. However, as mentioned above, LPA does not calculate any radiative coupling to and from guided modes by construction. Reference [15] derived a convergent series of perturbative corrections to the LUA, the limit of LPA as the period goes to zero, and used higher-order terms to compute coupling to guided modes in nearly uniform structures. However, that work has not been extended to handle locally periodic surfaces, requires the implementation of complicated integral-equation operators (far more expensive than the overlap integrals in this work), and does not exploit an explicit decomposition of the fields into guided and radiative modes. Other methods of domain decomposition besides LUA and LPA have been proposed to handle strong near-field intercell coupling, such as overlapping domains that model the overlapping regions from neighboring unit cells [61], combined with absorbing-wall domains that model each unit cell with perfectly matched layer boundary conditions [62]; however, these methods explicitly discard long-range interactions and hence omit guided-mode propagation. We also note that many techniques have been developed to design large-scale metasurface antennas, filters, and other devices at optical and microwave frequencies, but these make use of *leaky* waves rather than completely *guided* waves and thus solve a fundamentally different problem [36–44].

Our framework builds on previous work that derived a CWT for locally periodic surfaces using a continuously varying basis of Bloch modes, derived by “lifting” into a higher-dimensional space of phase-shifted surfaces and

projecting back down to the physical result at the end [4]. The end result was a set of coupled ordinary differential equations in the mode coefficients that is straightforward to apply: the coupling coefficients are given by simple modal-overlap integrals proportional to the rate of change of the unit cells, using modes computed from small *periodic* Maxwell solves, and the equations can be solved to any order in the rate of change. (One does not even need to solve for every unit cell in a long taper but instead can interpolate the overlaps from a sample of intermediate cells [4,56–60], achieving *better* than linear scaling in the system diameter.) However, that framework cannot be directly applied to a metasurface system in which the incident wave is a radiative mode illuminating the whole surface from the *side* (as opposed to incident guided modes from the *ends*), and moreover, a discrete modal-expansion framework is difficult to use with a continuum of radiative modes. To address these limitations, the key feature we introduce here is that the LPA (radiative) solution appears as a zeroth-order *source* term in CWT involving an overlap integral between the LPA and guided-mode fields. As a consequence, we can solve for the guided-mode coefficients to first order in the rate of change by simple modal-overlap integrals involving the LPA source term (Algorithm 1). Our framework handles plane waves with arbitrary incident angles, which affect the LPA solution and introduce a phase velocity into the overlap integral. Moreover, one can calculate the coupling from non-plane-wave sources by a plane-wave expansion (Fourier transform) of the source. Our derivation and validation in this paper are for metasurfaces in two spatial dimensions that vary only in one direction (z), but in the concluding remarks (Sec. IV) we discuss extension to surfaces in three spatial dimensions that vary in two directions: the key idea is that the variations in each direction decouple at first order and therefore can be treated separately.

II. LOCALLY UNIFORM FRAMEWORK

The problem setup is depicted in Fig. 1(a). We consider a surface and surrounding medium in two spatial dimensions that are characterized by the dielectric function $\varepsilon(x, z)$ and the magnetic permeability function $\mu(x, z)$. The medium surrounding the surface is homogeneous. We choose the surface to be locally uniform along the z direction; that is, the cross section along the x direction varies continuously and slowly with z . A plane wave of frequency ω from the surrounding medium is incident upon the surface with wave vector $\boldsymbol{\beta} = (\beta_x, \beta_0)$ (propagation constant β_0). Here, the goal is to solve for the coupling of the incident plane wave to guided waves in the surface.

A. Coupled-wave theory for locally uniform surfaces

To solve this problem, we apply the well-known spatial coupled-wave theory for locally uniform surfaces, following the approach and notation found in Secs. II and III of [4]. This section serves as a review of coupled-wave theory, for which more details can be found in [4].

By isolating the longitudinal (z) from the transverse (xy) derivatives, the fully vectorial source-free Maxwell's equa-

tions can be rewritten at a fixed frequency ω as

$$\hat{A}|\psi\rangle = -i\frac{\partial}{\partial z}\hat{B}|\psi\rangle, \quad (1)$$

where $|\psi\rangle$ is the four-component column vector

$$|\psi\rangle \equiv \begin{pmatrix} \mathbf{E}_{xy}(x, y, z) \\ \mathbf{H}_{xy}(x, y, z) \end{pmatrix} e^{-i\omega t} \quad (2)$$

involving the transverse (xy) electric and magnetic fields $\{\mathbf{E}_{xy}, \mathbf{H}_{xy}\}$ and \hat{A} and \hat{B} are the matrices

$$\hat{A} \equiv \begin{pmatrix} \omega\varepsilon/c - \frac{c}{\omega}\nabla_{xy} \times \frac{1}{\mu}\nabla_{xy} \times & 0 \\ 0 & \omega\mu/c - \frac{c}{\omega}\nabla_{xy} \times \frac{1}{\varepsilon}\nabla_{xy} \times \end{pmatrix}, \quad (3)$$

$$\hat{B} \equiv \begin{pmatrix} 0 & -\hat{\mathbf{z}} \times \\ \hat{\mathbf{z}} \times & 0 \end{pmatrix} = \begin{pmatrix} & & 1 \\ & -1 & \\ 1 & & \end{pmatrix}, \quad (4)$$

where $\nabla_{xy} = \frac{\partial}{\partial x}\hat{\mathbf{x}} + \frac{\partial}{\partial y}\hat{\mathbf{y}}$. The inner product of two column vectors $|\psi\rangle$ and $|\psi'\rangle$ is given by

$$\langle\psi|\psi'\rangle \equiv \int \mathbf{E}_{xy}^* \cdot \mathbf{E}'_{xy} + \mathbf{H}_{xy}^* \cdot \mathbf{H}'_{xy}, \quad (5)$$

where the integral is over the cross section along x at a fixed z . Under this inner product, \hat{A} and \hat{B} are Hermitian operators for real and lossless ε and μ .

Equation (1) can be solved using spatial coupled-wave theory, which employs an expansion basis of “instantaneous” uniform waveguides defined by the cross sections of the surface [see Fig. 1(a)]. The fields at each cross section are expanded in terms of the instantaneous-waveguide modes, resulting in a set of coupled differential equations in the mode coefficients. The instantaneous-waveguide modes are found by solving Maxwell's equations [Eq. (1)] for an infinite uniform waveguide. Since there is translational symmetry in the z direction, $|\psi\rangle$ can be chosen to satisfy $|\psi\rangle = |\beta\rangle e^{i\beta z}$, where β is the propagation constant and $|\beta\rangle$ is a z -independent function. Equation (1) becomes

$$\hat{A}|\beta\rangle = \beta\hat{B}|\beta\rangle. \quad (6)$$

Assuming real ε and μ , (6) is a generalized Hermitian eigenproblem with eigenvalue β . It follows that there is an orthogonality relation between the eigenstates: $\langle\beta|\hat{B}|\beta'\rangle = 0$ when $\beta \neq \beta'$, where it is assumed that β and β' are real (corresponding to propagating modes). If the eigenstates are guided modes, one can label the modes by an integer $n = 0, 1, 2, \dots$ since the eigenvalues are discrete, taking the state with eigenvalue β_n to be $|n\rangle$. The labeling can be extended to the continuum of nonguided modes as well since in practice the eigenvalues of such modes become discrete in a finite-size computational cell. Then, the propagating (real- β) modes can be normalized such that

$$\langle m|\hat{B}|n\rangle = \delta_{m,n}\eta_n, \quad (7)$$

where $\eta_n = \pm 1$ for forward and backward propagating modes. In coupled-wave theory, Eq. (6) must be solved for *each* cross section (z) of the locally uniform surface. The eigenproblem

is therefore labeled with z :

$$\hat{A}|n\rangle_z = \beta_n(z)\hat{B}|n\rangle_z, \quad (8)$$

where $|n\rangle_z$ and $\beta_n(z)$ are the instantaneous eigenmodes and eigenvalues and the z dependence of \hat{A} is implicit.

The fields of the locally uniform surface are then expanded in terms of the instantaneous eigenmodes by considering the following ansatz for $|\psi\rangle$ in Eq. (1):

$$|\psi(z)\rangle = \sum_n c_n(z)|n\rangle_z \exp\left(i \int^z \beta_n(z') dz'\right), \quad (9)$$

where $c_n(z)$ are eigenmode coefficients describing the intermodal scattering along z and the phase of each mode is given by an integral over $\beta_n(z)$. Next, one substitutes the ansatz into Eq. (1) and solves for the mode evolution in z . The result is a linear differential equation in the mode coefficients

$$\begin{aligned} \frac{dc_m}{dz} = & -\eta_m \sum_{n \neq m} \frac{\langle m | \frac{\partial \hat{A}}{\partial z} | n \rangle_z}{\beta_n(z) - \beta_m(z)} \\ & \times \exp\left(i \int^z [\beta_n(z') - \beta_m(z')] dz'\right) c_n \\ & - \eta_m \langle m | \hat{B} \frac{\partial |m\rangle_z}{\partial z} c_m \end{aligned} \quad (10)$$

that relates the mode evolution to the rate of change of the eigenoperator $\partial \hat{A} / \partial z$ and the eigenmode fields and propagation constants. Note that the right-hand side of Eq. (10) includes a ‘‘self-interaction’’ term given by $-\eta_m \langle m | \hat{B} \frac{\partial |m\rangle_z}{\partial z} c_m$. Appendix B of [4] shows that by imposing a simple phase choice on $|m\rangle_z$, corresponding to a Berry phase [63], this term can be *exactly* eliminated. If the instantaneous eigenmodes are chosen to be purely real, the condition reduces to choosing the overall sign of the modes consistently. Moving forward, we will thus omit the self-interaction term. Equation (10) is the ‘‘coupled-wave equations’’ of CWT and has been derived in a variety of notations by many authors [1–4]; its historical roots trace back to the ‘‘telegrapher’s equations’’ for transmission lines.

Equation (10) can be solved approximately by perturbatively expanding in the rate of change of the surface cross section. One assumes that $\partial \hat{A} / \partial z$ is small, i.e., the surface cross section varies slowly, and expands the mode coefficients in powers of $\partial \hat{A} / \partial z$:

$$c_m(z) = c_m^{(0)}(z) + c_m^{(1)}(z) + c_m^{(2)}(z) + \dots \quad (11)$$

Substituting the expansion into Eq. (10) and solving to first order in $\partial \hat{A} / \partial z$ yield

$$\begin{aligned} \frac{dc_m^{(1)}}{dz} = & -\eta_m \sum_{n \neq m} \frac{\langle m | \frac{\partial \hat{A}}{\partial z} | n \rangle_z}{\beta_n(z) - \beta_m(z)} \\ & \times \exp\left(i \int^z [\beta_n(z') - \beta_m(z')] dz'\right) c_n^{(0)}, \end{aligned} \quad (12)$$

which gives the first-order mode coefficients in terms of those at zeroth order.

B. Combining CWT with LUA

Physically, we can think of the zeroth-order terms as those that do not result from any intermodal scattering and instead couple directly to the incident wave. Those at first order are the result of zeroth-order terms that have undergone a single-scattering process from the nonzero rate of change. Let us therefore consider the zeroth-order terms in the context of an incident plane wave from the side [Fig. 1(a)]. A zeroth-order solution to the scattering problem neglects the rate of change of the surface, and hence, it conserves the propagation constant of the incident plane wave. Therefore, the zeroth-order modes consist only of radiation modes with propagation constant β_0 and no guided modes. We thus set $\beta_n(z) = \beta_0$ in Eq. (12). In principle, we can calculate the guided-mode couplings to first order in $\partial \hat{A}_z / \partial z$ by computing the zeroth-order radiation-mode fields and propagation constants (as well as the guided mode of interest). However, this procedure is not entirely convenient, as it requires normalizing radiation modes that lie in a continuous spectrum.

To simplify the approach, we employ LUA [15,18–20], which is equivalent to LPA in the limit of zero period. In LUA, the fields scattered from a locally uniform structure are approximated by a composition of scattering problems from uniform waveguides. That is, the fields at each cross section are obtained by scattering off of the completely uniform waveguide defined by the cross section [see Fig. 1(a)]. While LUA yields a z -dependent solution that captures the surface variation, it is a zeroth-order approximation because the incident field interacts only with z -independent structures. For an incident plane wave with propagation constant β_0 , the LUA solution is given by

$$|\psi_{\text{LUA}}\rangle = \sum_n c_n^{(0)}(z)|n\rangle_z, \quad (13)$$

where conservation of the wave vector implies that n sums over radiation modes of a fixed frequency *and* propagation constant (β_0) and the mode coefficients $c_n^{(0)}(z)$ are zeroth order in the rate of change. For convenience, we are factoring out the phase $e^{i\beta_0 z}$; that is, the full zeroth-order fields are given by $|\psi_{\text{LUA}}\rangle e^{i\beta_0 z}$.

It is clear from Eq. (13) that the LUA solution contains the zeroth-order solution to coupled-wave theory for an incident plane wave. After setting $\beta_n(z) = \beta_0$ in Eq. (12), pulling out the sum over n , and collecting the n -dependent terms, we can then insert the LUA solution:

$$\begin{aligned} c_m^{(1)}(L) = & -\eta_m \int_0^L dz \frac{\langle m | \frac{\partial \hat{A}}{\partial z} | \psi_{\text{LUA}} \rangle}{\beta_0 - \beta_m(z)} \\ & \times \exp\left(i \int^z [\beta_0 - \beta_m(z')] dz'\right), \end{aligned} \quad (14)$$

where we have integrated both sides of Eq. (12) to solve explicitly for the guided-mode coefficients. We can further simplify our result by writing the inner-product integral in terms of the guided-mode and LUA fields (Eq. (17) of [4]). For instance, for a locally uniform surface with constant μ

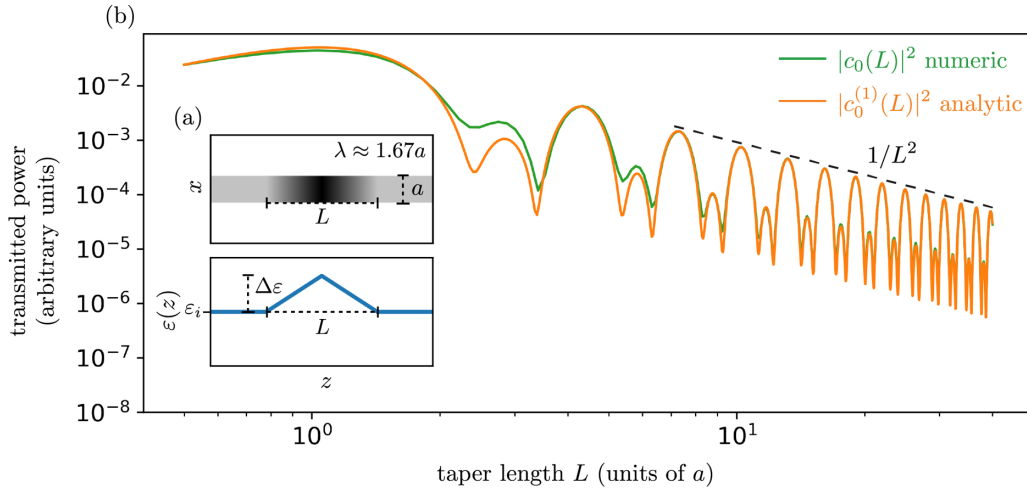


FIG. 2. (a) Top: locally uniform surface with thickness a under consideration, with permittivity indicated by the shading. Bottom: Plot of the permittivity $\varepsilon(z)$ as a function of z . Here $\varepsilon_i = 1.25$, and $\Delta\varepsilon = 0.5$. (b) Transmitted power of the fundamental TM guided mode at $\omega = 0.6 (2\pi c/a)$, or $\lambda \approx 1.67(a)$, as a function of taper length L . Results from brute-force FDTD simulations are shown in green, and results from our semianalytic framework [Eq. (15)] are shown in orange.

and z -varying ε , Eq. (14) becomes

$$c_m^{(1)}(L) = \frac{-\eta_m \omega}{c} \int_0^L dz \frac{\exp(i \int^z [\beta_0 - \beta_m(z')] dz')}{\beta_0 - \beta_m(z)} \times \frac{d\varepsilon}{dz} \int_z dx \mathbf{E}_m^* \cdot \mathbf{E}_{\text{LUA}}, \quad (15)$$

where \mathbf{E}_m and \mathbf{E}_{LUA} are the three-component electric fields of the guided-mode and LUA solutions, respectively, and the integral over the cross section includes only the region where $d\varepsilon/dz \neq 0$, i.e., between the boundaries of the waveguide.

C. Validation

Here, we validate our framework for locally uniform surfaces against brute-force FDTD simulations carried out in a freely available software package [64]. We consider a surface with locally uniform (slowly varying) permittivity $\varepsilon(z)$ depicted in Fig. 2: a triangular taper in $\varepsilon(z)$ joining two uniform waveguides. In this example, the uniform waveguides have permittivity $\varepsilon_i = 1.25$; the taper has an amplitude $\Delta\varepsilon = 0.5$ and a variable length L . The surface has a thickness a and is surrounded on either side by air ($\varepsilon = 1$). We consider a normally incident plane wave ($\beta_0 = 0$) with TM polarization (such that $\mathbf{E} = E\hat{\mathbf{y}}$ is perpendicular to the xz plane) coupling to the fundamental TM guided mode. The operating frequency is $0.6 (2\pi c/a)$, or $\lambda \approx 1.67(a)$.

The choice of basis for CWT and LPA in this case is straightforward; we simply choose a set of uniform waveguides defined by the cross sections of the taper (see Fig 2). For a particular waveguide in the basis, we compute the guided-mode fields and propagation constant using a freely available eigensolver of Maxwell's equations [65], and we compute the LPA fields (given the same normally incident plane wave) using a one-dimensional (1D) FDTD simulation for each cross section. We then use the guided mode and LPA solutions to compute the overlap integrals over the cross sections and the phase terms in Eq. (15). In principle, one needs to do

this for each cross section along the taper; however, because the overlap integral and propagation constants are continuous functions of z , it is sufficient to compute these quantities for only a few points and interpolate. Moreover, the interpolation can be reused for any taper length L by rescaling with the rate of change. We note that one must be careful to choose the phase of the guided-mode fields consistently across different points, which eliminates a “self-coupling” term from the derivation ([4], Appendix B).

Figure 2 shows the validation results: for a range of taper lengths L , we compute the transmitted power of the guided mode using both our framework (orange) and a brute-force FDTD simulation (green). The plot shows excellent agreement in the large- L (small rate of change) limit and relatively good agreement even for smaller L comparable to the wavelength. As expected from the slope discontinuity of $\Delta\varepsilon$, the coupled power decreases as $1/L^2$, which one can show analytically by Fourier analysis via a change of variables in CWT [57].

III. LOCALLY PERIODIC FRAMEWORK

In this section, we derive the coupling of an incident plane wave to guided waves on a locally *periodic* surface. The problem setup is depicted in Fig. 1(b). The setup is the same as in Sec. II, except that the surface is now locally periodic in the z direction rather than locally uniform; that is, the unit cells that compose the surface vary slowly with z . Since a locally periodic surface becomes locally uniform in the limit of zero period, the results of this section are a strict generalization of Sec. II. As is also discussed in [4], the framework developed in Sec. II would be a poor approximation for a photonic-crystal structure in which the periodic unit cell is rapidly varying on a subwavelength scale—it is critical to take the local periodicity into account analytically to obtain an accurate first-order approximation.

A. CWT for locally periodic surfaces

To solve this problem, we apply spatial coupled-wave theory for locally *periodic* surfaces, which was developed in Sec. IV of [4]. This section will largely review the coupled-wave theory developed in [4]; however, some modifications will be made due to the nature of the incident wave, which is then incorporated into the LPA+CWT formulation in the following section. As in Sec. II, one describes the fields scattered from the locally periodic surface by writing Maxwell's equations at a fixed frequency ω [Eqs. (1)–(5)]. However, in this case, one considers an expansion basis of eigenmodes of instantaneous *periodic* waveguides at each cross section. In order to describe the waveguides and their eigenstates, one needs to introduce a “virtual” coordinate \tilde{z} , to be distinguished from z , that describes the variation within the unit cell of each waveguide [see Fig. 1(b)].

One may be interested in modeling a surface where the periods of the unit cells vary with z . If the periods of the unit cells are given by $\Lambda(z)$, one can define dimensionless coordinates scaled by $\Lambda(z)$ so that the instantaneous waveguides all have unit period. One can define $\zeta \equiv \int^z dz'/\Lambda(z')$, which effectively counts the number of periods in the locally periodic surface up to some point z , and $\tilde{\zeta} \equiv \tilde{z}/\Lambda(z)$. In the following, we will consider the case of a fixed period $\Lambda(z) = \Lambda$ for simplicity, in which case $\zeta = z/\Lambda$ and $\tilde{\zeta} = \tilde{z}/\Lambda$. In the Appendix, we generalize to the case of a z -dependent period. For further simplicity, we assume herein that the locally periodic surface has z -dependent ε but constant μ .

At each z , one defines an instantaneous periodic waveguide by $\varepsilon_z(\tilde{\zeta})$, where $\varepsilon_z(\tilde{\zeta}) = \varepsilon_z(\tilde{\zeta} + 1)$; that is, the waveguide is unit periodic in $\tilde{\zeta}$ space. To connect the instantaneous waveguides to the physical locally periodic surface, one demands that

$$\varepsilon_z(\tilde{\zeta} = \zeta) = \varepsilon(z). \quad (16)$$

That is, the instantaneous waveguide matches the physical surface at a single cross section, where $\tilde{\zeta} = \zeta$ (or $\tilde{z} = z$ since the period here is fixed).

Given each $\varepsilon_z(\tilde{\zeta})$, one can solve for the corresponding waveguide modes. For an infinite periodic waveguide with no current sources, Maxwell's equations at a fixed frequency ω are given by

$$\hat{A}(\tilde{\zeta})|\psi\rangle = -\frac{i}{\Lambda} \frac{\partial}{\partial \tilde{\zeta}} \hat{B}|\psi\rangle, \quad (17)$$

where $\hat{A}(\tilde{\zeta})$ is \hat{A} from Eq. (3) with $\varepsilon = \varepsilon_z(\tilde{\zeta})$ (where the z dependence of the equation is suppressed for now). There is discrete translational symmetry in the $\tilde{z}/\tilde{\zeta}$ direction, so by Bloch's theorem one can choose $|\psi\rangle = |\beta\rangle e^{i\beta\tilde{z}} = |\beta\rangle e^{i\beta\Lambda\tilde{\zeta}}$, where $|\beta\rangle$ is a unit-periodic function in $\tilde{\zeta}$ space. Equation (17) becomes

$$\hat{C}(\tilde{\zeta})|\beta\rangle = \beta\hat{B}|\beta\rangle, \quad (18)$$

where $\hat{C}(\tilde{\zeta}) \equiv \hat{A}(\tilde{\zeta}) + \frac{i}{\Lambda} \frac{\partial}{\partial \tilde{\zeta}} \hat{B}$. As in Sec. II, this is a generalized Hermitian eigenproblem for real ε and μ , which implies an orthogonality relation between the eigenstates: $\langle\beta|\hat{B}|\beta'\rangle = 0$ when $\beta \neq \beta'$ (assuming β and β' are real, corresponding to propagating and nonevanescing modes). By Bloch's theorem, eigenstates separated by reciprocal lattice vectors are equivalent up to a phase, i.e., $|\beta + \frac{2\pi}{\Lambda}\ell\rangle = e^{(-2\pi i/\Lambda)\ell\tilde{z}}|\beta\rangle$ for any

integer ℓ . This implies an extended orthogonality relation:

$$\langle\beta|\hat{B}e^{(-2\pi i/\Lambda)\ell\tilde{z}}|\beta'\rangle = 0 \quad (19)$$

when $\beta \neq \beta' + (2\pi/\Lambda)\ell$. Below, we choose the modes to lie within the first Brillouin zone, i.e., $\beta \in (-\pi/\Lambda, \pi/\Lambda]$, and label the modes $|n\rangle$ corresponding to discrete eigenvalues β_n , normalized as in Eq. (7). Finally, since a different waveguide is defined for each cross section of the locally periodic surface, one labels the eigenequation with z . Equation (18) becomes

$$\hat{C}_z(\tilde{\zeta})|n(\tilde{\zeta})\rangle_z = \beta_n(z)\hat{B}|n(\tilde{\zeta})\rangle_z, \quad (20)$$

where $\hat{C}_z(\tilde{\zeta}) = \hat{A}_z(\tilde{\zeta}) + \frac{i}{\Lambda} \frac{\partial}{\partial \tilde{\zeta}} \hat{B}$ and the $\tilde{\zeta}$ dependence of the eigenstates is explicitly denoted.

To employ coupled-wave theory, one needs to turn the instantaneous eigenmodes $|n(\tilde{\zeta})\rangle_z$ into an expansion basis for $|\psi(z)\rangle$. However, the eigenmodes are defined over a unit cell, whereas one needs to expand in a single cross section. One might try to expand in the $|n(\tilde{\zeta} = \zeta)\rangle_z$ modal cross section, but the $\tilde{\zeta}$ dependence must be retained in order to employ the orthogonality relation between the eigenmodes [Eq. (7)]. To resolve this, the approach of [4] is to simultaneously solve a *family* of scattering problems involving different surfaces. In particular, consider the surface in z space defined by $\varepsilon_z(\zeta)$. By the matching condition in Eq. (16), $\varepsilon_z(\zeta) = \varepsilon(z)$; that is, this defines the locally periodic surface in the physical scattering problem. Now instead, consider the entirely different surface in z space defined by $\varepsilon_z(\zeta + \Delta\tilde{\zeta})$, which shifts the argument away from ζ by an amount $\Delta\tilde{\zeta}$. One can imagine constructing this surface by doing the following: at each z , take the periodic waveguide defined by $\varepsilon_z(\tilde{\zeta})$, and insert the cross section at $\tilde{\zeta} = \zeta + \Delta\tilde{\zeta}$. This defines an entire family of surfaces parameterized by $\Delta\tilde{\zeta}$, where $\Delta\tilde{\zeta} = 0$ corresponds to the physical surface of interest. Since the instantaneous waveguides are unit periodic in $\Delta\tilde{\zeta}$, so too is the family of surfaces; that is, $\Delta\tilde{\zeta}$ and $\Delta\tilde{\zeta} + 1$ define the same surface. Note that the notation here differs slightly from [4], in that we are referring to the shift from ζ as $\Delta\tilde{\zeta}$ rather than $\tilde{\zeta}$.

In order to apply this framework to a scattering problem with an incident plane wave in free space, our approach will differ from that of [4]. In particular, just as a family of different surfaces is defined, we also define different incident waves—that is, we solve a family of problems with different surfaces *and* incident waves. To generate the family of incident waves, our technique will mirror that of the surfaces described above.

Let us suppose that the electric field of the incident plane wave is given by $\mathbf{E}(z) = \mathbf{E}_0 e^{i(\beta_0 z + \beta_x x)}$, where \mathbf{E}_0 is a constant three-component vector and the x dependence of $\mathbf{E}(z)$ is implicit. Now suppose that for each instantaneous periodic waveguide in \tilde{z} space, we imagine sending in an instantaneous incident plane wave with the *same* wave vector. We will label the electric fields of such incident waves by $\mathbf{E}_z(\tilde{\zeta})$, where the z subscript denotes that there is a different incident wave for *each* instantaneous periodic waveguide [just as $\varepsilon_z(\tilde{\zeta})$ was defined]. In analogy to the $\varepsilon_z(\tilde{\zeta})$, we impose a constraint on $\mathbf{E}_z(\tilde{\zeta})$ to connect them to the physical incident wave:

$$\mathbf{E}_z(\tilde{\zeta} = \zeta) = \mathbf{E}(z); \quad (21)$$

that is, the instantaneous incident waves match the physical incident wave at the $\tilde{\zeta} = \zeta$ (or $\tilde{z} = z$) cross section. We also demand that $\mathbf{E}_z(\tilde{\zeta})$ are themselves plane waves with the same wave vector as the physical incident wave. Given these constraints, there is only one possible choice of instantaneous incident waves:

$$\mathbf{E}_z(\tilde{\zeta}) = \mathbf{E}_0 e^{i(\beta_0 \Lambda \tilde{\zeta} + \beta_x \tilde{x})} = \mathbf{E}_0 e^{i(\beta_0 \tilde{z} + \beta_x \tilde{x})}, \quad (22)$$

i.e., plane waves with the same phase as the physical plane wave. Since all of the instantaneous incident waves are the same, we no longer need the z subscript, but we will keep it in order to make the analogy to $\varepsilon_z(\tilde{\zeta})$ clear. (For the case of variable-period structures in the Appendix, the analogous approach leads to an incident wave that is effectively no longer a plane wave in the unit-cell problems but which is then expanded in plane waves.)

We can now generate the family of incident waves just as [4] did for the family of surfaces. For instance, consider the incident wave in z space defined by $\mathbf{E}_z(\zeta)$ —by the matching condition in Eq. (21), this defines the incident wave in the physical scattering problem. Now, consider the incident wave in “shifted” z space, defined by $\mathbf{E}_z(\zeta + \Delta\tilde{\zeta}) = \mathbf{E}_0 e^{i\beta_0 \Lambda \Delta\tilde{\zeta}} e^{i(\beta_0 z + \beta_x x)}$. This defines a family of incident waves parameterized by $\Delta\tilde{\zeta}$ that have the same wave vector but phase-shifted amplitudes given by $e^{i\beta_0 \Lambda \Delta\tilde{\zeta}}$. The physical incident wave of interest corresponds to $\Delta\tilde{\zeta} = 0$, i.e., an unshifted incident wave.

One can now rewrite Maxwell’s equations [Eq. (1)] to solve the family of systems parameterized by $\Delta\tilde{\zeta}$:

$$\hat{A}_z(\zeta + \Delta\tilde{\zeta})|\psi\rangle_{\Delta\tilde{\zeta}} = -i \frac{\partial}{\partial z} \hat{B}|\psi\rangle_{\Delta\tilde{\zeta}}, \quad (23)$$

where $\varepsilon_z(\zeta + \Delta\tilde{\zeta})$ has been inserted into Eq. (1) through $\hat{A}_z(\zeta + \Delta\tilde{\zeta})$ and the fields have been labeled with $\Delta\tilde{\zeta}$. The family of incident waves described above are taken as boundary conditions. For each system, one expands $|\psi\rangle_{\Delta\tilde{\zeta}}$ in the basis of $|n(\zeta + \Delta\tilde{\zeta})\rangle_z$, i.e., the instantaneous-waveguide modes evaluated at the $\zeta + \Delta\tilde{\zeta}$ cross section:

$$|\psi(z)\rangle_{\Delta\tilde{\zeta}} = \sum_n c_n(z, \Delta\tilde{\zeta}) |n(\zeta + \Delta\tilde{\zeta})\rangle_z \times \exp\left(i \int^z \beta_n(z') dz'\right), \quad (24)$$

where the mode coefficients c_n have both z and $\Delta\tilde{\zeta}$ dependence. One now considers what happens to the system when $\Delta\tilde{\zeta}$ is increased by 1, i.e., $\Delta\tilde{\zeta} \rightarrow \Delta\tilde{\zeta} + 1$. Since $\varepsilon_z(\tilde{\zeta})$ is unit periodic, the surface remains the same, while the incident wave remains the same up to a phase $e^{i\beta_0 \Lambda}$. Therefore, the fields $|\psi(z)\rangle_{\Delta\tilde{\zeta}}$ pick up the same phase. This is in contrast to [4], where the fields remained the same under $\Delta\tilde{\zeta} \rightarrow \Delta\tilde{\zeta} + 1$ (due to the incident wave being guided rather than a plane wave). Here, we choose to absorb this phase within the mode coefficients, such that $c_n(z, \Delta\tilde{\zeta} + 1) = c_n(z, \Delta\tilde{\zeta}) e^{i\beta_0 \Lambda}$. Since the coefficients are therefore Bloch periodic, we can write them as the product of a periodic function and a $\Delta\tilde{\zeta}$ -dependent phase:

$$c_n(z, \Delta\tilde{\zeta}) = e^{i\beta_0 \Lambda \Delta\tilde{\zeta}} \tilde{c}_n(z, \Delta\tilde{\zeta}), \quad (25)$$

where $\tilde{c}_n(z, \Delta\tilde{\zeta} + 1) = \tilde{c}_n(z, \Delta\tilde{\zeta})$. Following [4], since $\tilde{c}_n(z, \Delta\tilde{\zeta})$ are periodic in $\Delta\tilde{\zeta}$, one can expand \tilde{c}_n as a Fourier series:

$$\tilde{c}_n(z, \Delta\tilde{\zeta}) = \sum_{\ell} \tilde{c}_{n,\ell}(z) e^{2\pi i \ell \Delta\tilde{\zeta}}. \quad (26)$$

Using this expression, we will formulate the coupled-wave equations in $\tilde{c}_{n,\ell}$ rather than \tilde{c}_n or c_n . After solving explicitly for $\tilde{c}_{n,\ell}$, one can always recover the physical mode coefficients by the relation

$$c_n(z) = \tilde{c}_n(z) = \sum_{\ell} \tilde{c}_{n,\ell}(z), \quad (27)$$

i.e., by setting $\Delta\tilde{\zeta} = 0$.

Next, one substitutes the ansatz for $|\psi(z)\rangle_{\Delta\tilde{\zeta}}$ into Maxwell’s equations for $\varepsilon_z(\zeta + \Delta\tilde{\zeta})$ [Eq. (23)]. The result is a set of differential equations in $\tilde{c}_{n,\ell}$ that is identical to Eq. (25) of [4]:

$$\begin{aligned} \frac{d\tilde{c}_{m,k}}{dz} = & -\eta_m \sum_{n,\ell \neq m,k} \frac{\langle m | e^{2\pi i k \tilde{\zeta}} \frac{\partial \hat{C}_z}{\partial z} e^{-2\pi i \ell \tilde{\zeta}} | n \rangle_z}{\Delta\beta_{n,l;m,k}(z)} \\ & \times \exp\left(i \int^z \Delta\beta_{n,l;m,k}(z') dz'\right) \tilde{c}_{n,\ell} \\ & - \eta_m^* \langle m | \hat{B} \frac{\partial |m\rangle_z}{\partial z} \tilde{c}_{m,k}, \end{aligned} \quad (28)$$

where the phase mismatch $\Delta\beta$ is given by

$$\Delta\beta_{n,\ell;m,k}(z) \equiv \beta_n(z) - \beta_m(z) + \frac{2\pi}{\Lambda}(\ell - k). \quad (29)$$

As in Sec. II, the self-interaction term $-\eta_m^* \langle m | \hat{B} \frac{\partial |m\rangle_z}{\partial z} c_{m,k}$ can be *exactly* eliminated by a straightforward “Berry” phase choice for $|m\rangle_z$ ([4], Appendix B). This involves evaluating simple overlap integrals between instantaneous eigenmodes at nearby points z . We thus omit this term below. One can approximately solve Eq. (28) by perturbatively expanding the mode coefficients in the rate of change. That is, one assumes that $\partial \hat{C}_z(\tilde{\zeta})/\partial z$ is small, i.e., the instantaneous unit cell varies slowly with z , and expands $\tilde{c}_{m,k}$ in powers of $\partial \hat{C}_z(\tilde{\zeta})/\partial z$:

$$\tilde{c}_{m,k}(z) = \tilde{c}_{m,k}^{(0)}(z) + \tilde{c}_{m,k}^{(1)}(z) + \tilde{c}_{m,k}^{(2)}(z) + \dots \quad (30)$$

Substituting the expansion into Eq. (28) and solving to first order, one finds

$$\begin{aligned} \frac{d\tilde{c}_{m,k}^{(1)}}{dz} = & -\eta_m \sum_{n,\ell \neq m,k} \frac{\langle m | e^{2\pi i k \tilde{\zeta}} \frac{\partial \hat{C}_z}{\partial z} e^{-2\pi i \ell \tilde{\zeta}} | n \rangle_z}{\Delta\beta_{n,l;m,k}(z)} \\ & \times \exp\left(i \int^z \Delta\beta_{n,l;m,k}(z') dz'\right) \tilde{c}_{n,\ell}^{(0)}, \end{aligned} \quad (31)$$

which gives the first-order mode coefficients in terms of those at zeroth order (obtained below from the LPA solution).

B. Combining CWT with LPA

Here, we can make a similar argument to that in Sec. II regarding the zeroth-order modes. Since the zeroth-order solution necessarily neglects the rate of change of the unit cell, it must conserve the wave vector of the incident plane wave up to $2\pi/\Lambda$. Therefore, the zeroth-order modes are radiation modes with $\beta_n(z) = \beta_0$ (since we have arbitrarily restricted

the modes to lie within the first Brillouin zone). Then Eq. (31) becomes

$$\frac{d\tilde{c}_{m,k}^{(1)}}{dz} = -\eta_m \sum_{n,\ell \neq m,k} \frac{\langle m | e^{2\pi i k \tilde{\zeta}} \frac{\partial \tilde{c}_z}{\partial z} e^{-2\pi i \ell \tilde{\zeta}} | n \rangle_z}{\Delta\beta_{l,m,k}(z)} \times \exp\left(i \int^z \Delta\beta_{l,m,k}(z') dz'\right) \tilde{c}_{n,\ell}^{(0)}, \quad (32)$$

where the phase mismatch $\Delta\beta$ is now given by

$$\Delta\beta_{l,m,k}(z) \equiv \beta_0 - \beta_m(z) + \frac{2\pi}{\Lambda}(\ell - k). \quad (33)$$

We have found an expression for guided-mode coefficients to first order in $\partial\tilde{C}_z(\tilde{\zeta})/\partial z$. However, in its current form, we must solve explicitly for and normalize the zeroth-order radiation modes. To avoid this, we will express Eq. (32) in terms of the fields given by LPA, which is a generalization of LUA in Sec. II B to structures with nonzero periods.

In LPA, the fields scattered from a locally periodic surface are approximated by a composition of scattering problems from periodic waveguides [5–17,21]. In this case, it is natural to apply LPA alongside coupled-wave theory since we have already defined instantaneous periodic waveguides and incident waves corresponding to each cross section.

Suppose we want to apply LPA to the physical surface of interest. For a given cross section, we simply take the instantaneous waveguide defined by $\varepsilon_z(\tilde{\zeta})$ and send in the instantaneous incident wave with electric field $\mathbf{E}_z(\tilde{\zeta})$. For each z , this produces a solution that is defined over a unit period in $\tilde{\zeta}$ space. Since the incident wave is a plane wave with propagation constant β_0 , conservation of the wave vector implies that we can write the solution in $\tilde{\zeta}$ space as a sum over radiation modes with fixed frequency and propagation constant (β_0):

$$|\psi_{\text{LPA}}(\tilde{\zeta})\rangle_z = \sum_n b_n^{(0)}(z) |n(\tilde{\zeta})\rangle_z, \quad (34)$$

where the radiation modes are weighted by coefficients $b_n^{(0)}(z)$ that depend on z but not $\tilde{\zeta}$. As denoted by the superscript, the coefficients are zeroth order in $\partial\tilde{C}_z(\tilde{\zeta})/\partial z$ because the incident waves scatter off of completely periodic structures. For convenience, we are factoring out the phase $e^{i\beta_0 z}$ from the LPA fields; that is, the full fields are given by $|\psi_{\text{LPA}}(\tilde{\zeta})\rangle_z e^{i\beta_0 z}$. Hence, when we phase shift the unit cells by $\Delta\tilde{\zeta}$, the corresponding LPA fields for the shifted problem are $|\psi_{\text{LPA}}(\tilde{\zeta} + \Delta\tilde{\zeta})\rangle_z e^{i\beta_0 \Lambda(\tilde{\zeta} + \Delta\tilde{\zeta})}$.

Next, we need to understand how $b_n^{(0)}$ in (34) relate to the Fourier coefficients $\tilde{c}_{n,\ell}^{(0)}$ so that we can insert $|\psi_{\text{LPA}}\rangle$ into Eq. (32). First, if we compare the expression for the full LPA fields (including the $\Delta\tilde{\zeta}$ -dependent phase) to the modal expansion of Eq. (24) assuming that $\beta_n(z) = \beta_0$, we find that $b_n^{(0)}$ are equivalent to $\tilde{c}_n^{(0)}$; that is, they are periodic in $\Delta\tilde{\zeta}$. Moreover, by Eq. (34), $b_n^{(0)}$ are independent of $\tilde{\zeta}$ and therefore of $\Delta\tilde{\zeta}$. Hence, by Eq. (26), $b_n^{(0)}(z)$ are equivalent to the $\ell = 0$ Fourier coefficients $\tilde{c}_{n,0}^{(0)}$. Since the LPA solution is the zeroth-order solution to coupled-wave theory for an incident plane wave, we can drop the sum over ℓ in Eq. (32), leaving only the $\ell = 0$ term. Then, after pulling out the sum over n and collecting the n -dependent terms, we can insert the LPA

solution into Eq. (32):

$$\tilde{c}_m^{(1)}(L) = -\eta_m \int_0^L dz \sum_k \frac{\langle m | e^{2\pi i k \tilde{\zeta}} \frac{\partial \tilde{C}_z}{\partial z} | \psi_{\text{LPA}} \rangle_z}{\Delta\beta_{0,m,k}(z)} \times \exp\left(i \int^z \Delta\beta_{0,m,k}(z') dz'\right), \quad (35)$$

where we have integrated both sides over z and summed over k to solve explicitly for $\tilde{c}_m^{(1)}$ [Eq. (27)]. Finally, we can simplify this result by expressing the inner-product integral in terms of the guided-mode and LPA fields (Eq. (30) of [4]). Assuming that the locally periodic surface has z -independent μ , Eq. (35) becomes

$$\tilde{c}_m^{(1)}(L) = \frac{-\eta_m \omega}{c} \int_0^L dz \sum_k \frac{\exp\left(i \int^z \Delta\beta_{0,m,k}(z') dz'\right)}{\Delta\beta_{0,m,k}(z)} \times \int_z d\tilde{x} d\tilde{z} e^{(2\pi i/\Lambda)k\tilde{z}} \frac{\partial \varepsilon_z(\tilde{\zeta})}{\partial z} \mathbf{E}_m^* \cdot \mathbf{E}_{\text{LPA}}, \quad (36)$$

where \mathbf{E}_m and \mathbf{E}_{LPA} are the three-component electric fields of the guided-mode and LPA solutions, respectively. Equation (36) also assumes that $\partial\varepsilon_z(\tilde{\zeta})/\partial z$ has no moving boundary—to handle such cases, the integral must be written in terms of the field components that are continuous across the boundary, \mathbf{E}_{\parallel} and D_{\perp} ([66], Eq. (12)).

C. Validation

Here, we validate our framework for locally periodic surfaces against brute-force FDTD simulations. We also provide Algorithm 1, which describes how to efficiently apply our framework to compute the coupling to guided waves in such surfaces. We consider a surface that is analogous to the example in Sec. II C: the permittivity is described by a sine wave with a fixed period and a triangular taper in the amplitude, with uniform waveguides on either side of the taper. That is, the permittivity is given by $\varepsilon(z) = \varepsilon_i + \Delta\varepsilon A(z) \sin(2\pi z/\Lambda)$, where $A(z)$ describes a triangular taper with unit amplitude and width L (see Fig. 3). In this case, we choose $\varepsilon_i = 2$ (larger than in Sec. II C in order to ensure a strongly localized guided mode), $\Delta\varepsilon = 0.2$, and $\Lambda = 0.25a$. The surface has a thickness of a and is surrounded on either side by air. We consider a normally incident plane wave ($\beta_0 = 0$) with TM polarization coupling to a fundamental TM guided mode, operating at a frequency of $1.1(2\pi c/a)$.

In order to choose a basis of waveguides, we must obey the matching condition of Eq. (16). While there are many choices that satisfy this condition, the most straightforward one is as follows: for each z , the permittivity is given by a sine wave (with period Λ) whose amplitude matches that of the physical taper at z . That is, the instantaneous waveguides are described by $\varepsilon_z(\tilde{z}) = \varepsilon_i + \Delta\varepsilon A(z) \sin(2\pi \tilde{z}/\Lambda)$. As described in Sec. VI of [4], one must be careful to choose a taper where the guided mode remains guided throughout the taper, e.g., where the operating frequency does not fall into a band gap at some intermediate points—in such cases, the theory breaks down, and the transmission of the mode becomes poor. As in Sec. II C, for a particular instantaneous waveguide, we compute the guided mode and LPA solutions, then use these to compute the overlap integrals and phase terms in Eq. (36).

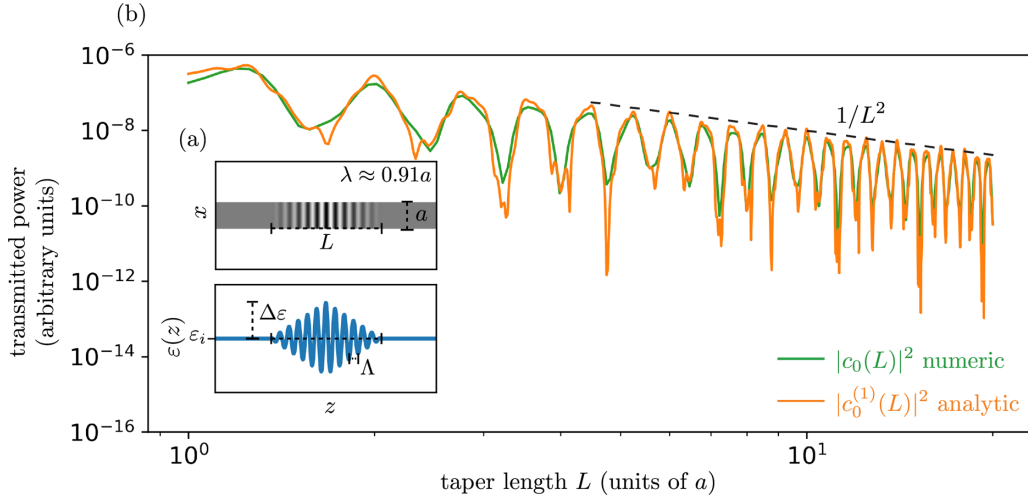


FIG. 3. (a) Top: locally periodic surface with thickness a under consideration. Bottom: Plot of the permittivity $\varepsilon(z)$ as a function of z , a sine wave with fixed period and a triangular taper in the amplitude. Here $\varepsilon_i = 2$, $\Delta\varepsilon = 0.2$, and $\Lambda = 0.25(a)$. (b) Transmitted power of the fundamental TM guided mode at $\omega = 1.1 (2\pi c/a)$, or $\lambda \approx 0.91(a)$, as a function of taper length L . Results from brute-force FDTD simulations are shown in green, and results from our semianalytic framework [Eq. (36)] are shown in orange.

While Eq. (36) sums over infinitely many integer values of k , in practice one can drop most of these terms. This is because the largest contributions come from terms where $\Delta\beta_{0,m,k}(z)$ is smallest due to the integral over the phase in Eq. (36). In this case, $\Delta\beta_{0,m,k}(z)$ may be smallest when $k = -1, 0$, or 1 depending on the propagation directions of the incident wave and guided mode (if both propagate in the same direction, it is when $k = 0$). We found that it is sufficient to sum over the five terms centered around the smallest $\Delta\beta_{0,m,k}(z)$; thus, using the seven terms from $k = -3$ to 3 should always be more than enough. (In this particular example, the average relative error between five and seven terms was 0.0037%.) As in Sec. II C, rather than computing the overlap integrals and propagation constants for each point along the taper, we do so only for a few points and interpolate the remainder; in this case, we perform separate interpolations of the overlap integral for each value of k . After doing so, the interpolations can be applied to any taper length L by rescaling with the rate of change. As before, one must consistently choose the phase of the guided-mode fields in order to disregard the self-interaction term omitted in the derivation.

Figure 3 shows the validation results: for each taper length L , we compute the transmitted power of the guided mode using both our framework (orange) and an FDTD simulation (green), which show agreement in the large- L limit. In the following, we provide an algorithm that summarizes the steps

involved in computing the coupling to a guided mode using our framework, assuming a surface whose unit cells have a fixed period, $\Lambda(z) = \Lambda$. In order to do so, we will rewrite Eq. (35) in a form that is more convenient to apply. First, we define a dimensionless coordinate $u = z/L$ such that the surface varies only from $u = 0$ to $u = 1$. Next, following [56,57,59], we define a dimensionless parameter $s \in [0, 1]$ that characterizes the instantaneous unit cells of the surface. For instance, for a dielectric waveguide composed of a sequence of periodic flanges, s could be proportional to the flange height (see Fig. 1 of [57]). Then, a given surface taper is described by the continuous function $s(u)$, which defines a sequence of instantaneous unit cells for each point of the taper. In general, the taper may be described by multiple independent parameters $s_n(u)$; however, we assume only one parameter for simplicity. Then, Eq. (35) can be rewritten as

$$\tilde{c}_m^{(1)}(L) = -\eta_m \int_0^1 du s'(u) \sum_k \frac{M_k[s(u)]}{\Delta\beta_k[s(u)]} \times \exp\left(i \int^u \Delta\beta_k[s(u')] du'\right), \quad (37)$$

where $M_k[s(u)] = \langle m | e^{2\pi i k \tilde{z}} \frac{\partial \hat{C}_z}{\partial z} | \psi_{\text{LPA}} \rangle_z$ and $\Delta\beta_k[s(u)] = \Delta\beta_{0,m,k}(z)$ are the k -dependent matrix elements and phase mismatches, respectively. Using this notation, the algorithm

Algorithm 1. Semianalytical radiative- and guided-mode coupling for slowly varying periodic surfaces.

-
- 1: **for** a set of points $s \in [0, 1]$
 - 2: Compute the LPA solution $|\psi_{\text{LPA}}\rangle_s$: periodic scattering solve
 - 3: Compute the phase-corrected, normalized eigenmode $|m\rangle_s$ and $\beta_m(s)$: periodic eigensolve
 - 4: **for** $k = -3$ to 3
 - 5: Compute overlap integrals $M_k[s(u)]$ and phase mismatches $\Delta\beta_k[s(u)]$ [Eq. (37)]
 - 6: **end for**
 - 7: **end for**
 - 8: **for** each k : Form interpolating polynomials or splines for $M_k[s(u)]$ and $\Delta\beta_k[s(u)]$ to obtain values at any s
-

for computing the coupling to a guided wave in a nearly periodic surface with fixed period, $\Lambda(z) = \Lambda$, is described in Algorithm 1.

Step 1 could employ a variety of interpolation algorithms, such as cubic splines or Chebyshev polynomials, to obtain high accuracy from a small number of samples. Once one has carried out the algorithm for a particular taper with length L and rate profile $s(u)$, computing the coupling for additional tapers requires minimal effort. In particular, one simply redefines $s(u) \in [0, 1]$ and L and executes steps 8 since the calculations used to sample the interpolating points in steps 1–7 can be reused. This method was used in [56,57] to optimize a taper profile for guided-mode transmission, which requires evaluating CWT for many different $s(u)$. Finally, as described in the Appendix, this algorithm can be applied to structures with z -dependent periods by Fourier transforming the incident wave in ζ and applying the algorithm to each plane-wave component.

IV. CONCLUDING REMARKS

We have developed a semianalytical framework combining LPA and CWT that solves the scattering problem of incident radiation coupling to guided waves in large-area, slowly varying surfaces. Our framework is particularly relevant to optical metasurfaces, where solving this problem by brute-force simulation is formidable due to their large diameters and where existing analytical approximations have largely focused on the problem of incident radiation coupling to outgoing radiation or leaky modes. A possible application of this work is to metasurface design through numerical optimization, building on previous optimization of metasurfaces by LPA [14,17,50–52] and guided-mode couplers using CWT [56,57,59,60], such as for large-area grating couplers, for nonlinear processes and lasing via long-lifetime trapping in guided modes, and for other devices exploiting nonlocal effects due to transport by guided waves in large surfaces.

Our derivation could be straightforwardly generalized in a variety of ways, for example, to other linear materials, such as anisotropic, magnetic, absorptive (complex β), or even magneto-optic or chiral media. Arbitrary incident waves can be treated by expanding them in plane waves, applying our framework to each plane-wave component. Coupling radiative modes to guided modes, as in the examples above, is exactly equivalent to coupling guided modes to radiation under a reciprocity or time-reversal transformation [67]. Surfaces with a z -varying periodicity can be handled by a coordinate transformation as in [4] (Appendix A). Although the examples in this paper were for 1D patterned surfaces in two dimensions, exactly the same equations are applicable in three dimensions for structures that are nearly periodic along *one* direction, such as waveguide tapers [59].

A more subtle generalization would be to two-dimensional patterned surfaces in three dimensions that are nearly periodic along *two* directions (xy), as in most practical metasurface devices. One promising approach is to perturbatively decouple the two directions: for variation in two directions, one can Taylor expand the mode coefficients in both $\partial\hat{C}_x(\tilde{\zeta})/\partial x$ and $\partial\hat{C}_y(\tilde{\zeta})/\partial y$ (the slow rates of change in x and y). To first

order in these rates, the variations in each direction decouple: the mode coefficient is given by a term that is first order in $\partial\hat{C}_x(\tilde{\zeta})/\partial x$ plus another term that is first order in $\partial\hat{C}_y(\tilde{\zeta})/\partial y$. These terms could be computed independently by applying Eq. (35) to the variation in each direction (while approximating the surface as completely periodic in the perpendicular direction for each “line” of unit cells). (A practical obstacle to research on similar semianalytical methods in three dimensions is simply that brute-force validation becomes vastly more expensive. This can be combated by carefully choosing the three-dimensional test problem, such as an array of *disconnected* high-symmetry scatterers amenable to fast multipole or integral-equation methods [68–70], with appropriate absorbing-boundary techniques for the guided modes [71–73].)

Moreover, our work could serve as the starting point to compute *next*-order scattering processes, such as the second-order coupling of incident to outgoing radiation involving an intermediate scattering to guided modes (e.g., to capture nonlinear effects experienced by the guided modes), or simply as a validation check for the first-order calculation.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Army Research Office through the Institute for Soldier Nanotechnologies at MIT under Awards No. W911NF-18-2-0048 and No. W911NF-13-D-0001 and by the Simons Foundation collaboration on Extreme Wave Phenomena. R.P. is also supported in part by MIT-IBM Watson AI Laboratory (Challenge No. 2415).

APPENDIX: EXTENSION TO SURFACES WITH z -DEPENDENT PERIOD $\Lambda(z)$

In Sec. III, we considered only surfaces where the period of the instantaneous unit cell is z independent, i.e., $\Lambda(z) = \Lambda$. Here, we extend our framework to handle a z -dependent period, which will reduce to solving multiple problems with z -independent periods. Recall that, in anticipation of this case, we introduced dimensionless coordinates scaled by $\Lambda(z)$. In particular, we defined $\zeta \equiv \int^z dz'/\Lambda(z')$, which counts the number of periods in the locally periodic surface up to some point z , and $\tilde{\zeta} \equiv \tilde{z}/\Lambda(z)$, which scales all of the instantaneous gratings to have unit period. Using these coordinate systems and assuming a z -dependent period, all of the analysis carried out in Eqs. (16)–(20) still holds. However, a problem arises when we try to define the instantaneous incident waves [Eqs. (21) and (22)]. That is, for general $\Lambda(z)$ we cannot define $\mathbf{E}_z(\tilde{\zeta})$ that are *plane waves* while satisfying the constraint of Eq. (21). As a result, the family of incident waves parameterized by $\Delta\tilde{\zeta}$ will no longer be plane waves phase shifted by $e^{i\beta_0\Lambda\Delta\tilde{\zeta}}$ and, in general, will not be plane waves at all. Thus, the approach requires some modification.

Intuitively, the problem stems from trying to express the incident plane wave in terms of the scaled coordinates. In particular, the physical incident wave [with electric field $\mathbf{E}(\zeta)$] is *no longer a plane wave* in ζ space since it has effectively been

rescaled by $\Lambda(z)$ into an arbitrary incident wave. Using scaled coordinates transforms the original scattering problem with an incident plane wave and z -dependent period into one with an arbitrary incident wave and z -independent period since $\varepsilon_z(\tilde{\zeta})$ are all unit periodic in $\tilde{\zeta}$. Therefore, we can treat this problem using the framework that we developed for surfaces with a z -independent period. In particular, our solution is to Fourier transform the incident wave in ζ space into a plane-wave basis and apply Eq. (35) to each plane-wave component. To get the total coupling to guided modes, the couplings from each plane wave can simply be added together due to the linearity of Maxwell's equations. For instance, suppose the electric fields

of the physical incident wave are given by

$$\mathbf{E}(\zeta) = \int d\tilde{\beta}_0 d\tilde{\beta}_x \alpha(\tilde{\beta}_0, \tilde{\beta}_x) e^{i(\tilde{\beta}_0 \zeta + \tilde{\beta}_x x)}, \quad (\text{A1})$$

where $\alpha(\tilde{\beta}_0, \tilde{\beta}_x)$ is given by the Fourier transform of $\mathbf{E}(\zeta)$. Then, for each nonzero $\alpha(\tilde{\beta}_0, \tilde{\beta}_x)$ such that $\tilde{\beta}_x < 0$, i.e., the plane wave is traveling towards the surface rather than away from it, one computes the coupling from the plane wave by applying Eq. (35) assuming $\beta_0 = \tilde{\beta}_0$ and using the set of scaled instantaneous periodic waveguides with unit period [$\Lambda(z) = 1$].

- [1] D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic Press, Holmdel, New Jersey, 2012).
- [2] B. Z. Katsenelenbaum, L. M. del Río, M. Pereyaslavets, M. S. Ayza, and M. Thumm, *Theory of Nonuniform Waveguides: The Cross-section Method* (Institution of Electrical Engineers, London, UK, 1998).
- [3] A.W. Snyder and J. Love, *Optical Waveguide Theory* (Springer Science & Business Media, Berlin, Germany, 1983).
- [4] S. G. Johnson, P. Bienstman, M. A. Skorobogatiy, M. Ibanescu, E. Lidorikis, and J. D. Joannopoulos, Adiabatic theorem and continuous coupled-mode theory for efficient taper transitions in photonic crystals, *Phys. Rev. E* **66**, 066608 (2002).
- [5] F. Aieta, P. Genevet, M. A. Kats, N. Yu, R. Blanchard, Z. Gaburro, and F. Capasso, Aberration-free ultrathin flat lenses and axicons at telecom wavelengths based on plasmonic metasurfaces, *Nano Lett.* **12**, 4932 (2012).
- [6] A. Arbabi, E. Arbabi, Y. Horie, S. M. Kamali, and A. Faraon, Planar metasurface retroreflector, *Nat. Photonics* **11**, 415 (2017).
- [7] M. Khorasaninejad, W. T. Chen, A. Y. Zhu, J. Oh, R. C. Devlin, C. Roques-Carmes, I. Mishra, and F. Capasso, Visible wavelength planar metalenses based on titanium dioxide, *IEEE J. Sel. Top. Quantum Electron.* **23**, 43 (2017).
- [8] F. Aieta, M. A. Kats, P. Genevet, and F. Capasso, Multiwavelength achromatic metasurfaces by dispersive phase compensation, *Science* **347**, 1342 (2015).
- [9] M. Khorasaninejad, F. Aieta, P. Kanhaiya, M. A. Kats, P. Genevet, D. Rousso, and F. Capasso, Achromatic metasurface lens at telecommunication wavelengths, *Nano Lett.* **15**, 5358 (2015).
- [10] M. Khorasaninejad, Z. Shi, A. Y. Zhu, W. T. Chen, V. Sanjeev, A. Zaidi, and F. Capasso, Achromatic metalens over 60 nm bandwidth in the visible and metalens with reverse chromatic dispersion, *Nano Lett.* **17**, 1819 (2017).
- [11] E. Arbabi, A. Arbabi, S. M. Kamali, Y. Horie, and A. Faraon, Controlling the sign of chromatic dispersion in diffractive optics with dielectric metasurfaces, *Optica* **4**, 625 (2017).
- [12] V.-C. Su, C. H. Chu, G. Sun, and D. P. Tsai, Advances in optical metasurfaces: Fabrication and applications, *Opt. Express* **26**, 13148 (2018).
- [13] B. Groever, C. Roques-Carmes, S. J. Byrnes, and F. Capasso, Substrate aberration and correction for meta-lens imaging: An analytical approach, *Appl. Opt.* **57**, 2973 (2018).
- [14] R. Pestourie, C. Pérez-Arancibia, Z. Lin, W. Shin, F. Capasso, and S. G. Johnson, Inverse design of large-area metasurfaces, *Opt. Express* **26**, 33732 (2018).
- [15] C. Pérez-Arancibia, R. Pestourie, and S. G. Johnson, Sideways adiabaticity: Beyond ray optics for slowly varying metasurfaces, *Opt. Express* **26**, 30202 (2018).
- [16] L. Verslegers, P. B. Catrysse, Z. Yu, W. Shin, Z. Ruan, and S. Fan, Phase front design with metallic pillar arrays, *Opt. Lett.* **35**, 844 (2010).
- [17] J. Cheng, S. Inampudi, and H. Mosallaei, Optimization-based dielectric metasurfaces for angle-selective multifunctional beam deflection, *Sci. Rep.* **7**, 12228 (2017).
- [18] M. Born, E. Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. (Cambridge University Press, Cambridge, 1999).
- [19] D. C. O'Shea, T. J. Suleski, A. D. Kathman, and D. W. Prather, *Diffractive Optics: Design, Fabrication, and Test* (SPIE Publications, Bellingham, WA, 2003).
- [20] A. G. Voronovich, *Wave Scattering from Rough Surfaces* (Springer, Berlin, 1994).
- [21] N. Yu and F. Capasso, Flat optics with designer metasurfaces, *Nat. Mater.* **13**, 139 (2014).
- [22] Z. Lin, C. Roques-Carmes, R. E. Christiansen, M. Soljačić, and S. G. Johnson, Computational inverse design for ultra-compact single-piece metalenses free of chromatic and angular aberration, *Appl. Phys. Lett.* **118**, 041104 (2021).
- [23] M. Mansouree, A. McClung, S. Samudrala, and A. Arbabi, Large-scale parametrized metasurface design using adjoint optimization, *ACS Photonics* **8**, 455 (2021).
- [24] R. E. Christiansen, Z. Lin, C. Roques-Carmes, Y. Salamin, S. E. Kooi, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, Fullwave Maxwell inverse design of axisymmetric, tunable, and multi-scale multi-wavelength metalenses, *Opt. Express* **28**, 33854 (2020).
- [25] W. T. Chen, A. Y. Zhu, and F. Capasso, Flat optics with dispersion-engineered metasurfaces, *Nat. Rev. Mater.* **5**, 604 (2020).
- [26] E. Kuester, M. Mohamed, M. Piket-May, and C. Holloway, Averaged transition conditions for electromagnetic fields at a metafilm, *IEEE Trans. Antennas Propag.* **51**, 2641 (2003).
- [27] C. L. Holloway, E. F. Kuester, J. A. Gordon, J. O'Hara, J. Booth, and D. R. Smith, An overview of the theory and applications of metasurfaces: The two-dimensional equivalents of metamaterials, *IEEE Antennas Propag. Mag.* **54**, 10 (2012).
- [28] C. L. Holloway, A. Dienstfrey, E. F. Kuester, J. F. O'Hara, A. K. Azad, and A. J. Taylor, A discussion on the interpretation and characterization of metafilms/metalsurfaces: The

- two-dimensional equivalent of metamaterials, *Metamaterials* **3**, 100 (2009).
- [29] S. A. Tretyakov, Metasurfaces for general transformations of electromagnetic fields, *Philos. Trans. A* **373**, 20140362 (2015).
- [30] A. Epstein and G. V. Eleftheriades, Passive lossless Huygens metasurfaces for conversion of arbitrary source field to directive radiation, *IEEE Trans. Antennas Propag.* **62**, 5680 (2014).
- [31] A. Epstein and G. V. Eleftheriades, Floquet-Bloch analysis of refracting Huygens metasurfaces, *Phys. Rev. B* **90**, 235127 (2014).
- [32] S. Jahani and Z. Jacob, All-dielectric metamaterials, *Nat. Nanotechnol.* **11**, 23 (2016).
- [33] N. Yu, P. Genevet, F. Aieta, M. A. Kats, R. Blanchard, G. Aoust, J.-P. Tetienne, Z. Gaburro, and F. Capasso, Flat optics: Controlling wavefronts with optical antenna metasurfaces, *IEEE J. Sel. Top. Quantum Electron.* **19**, 4700423 (2013).
- [34] M. Khorasaninejad and F. Capasso, Metalenses: Versatile multifunctional photonic components, *Science* **358**, eaam8100 (2017).
- [35] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University Press, Princeton, NJ, 2008).
- [36] M. Benzaouia, J. D. Joannopoulos, S. G. Johnson, and A. Karalis, Quasi-normal mode theory of the scattering matrix, enforcing fundamental constraints for truncated expansions, *Phys. Rev. Research* **3**, 033228 (2021).
- [37] G. Minatti, F. Caminita, E. Martini, M. Sabbadini, and S. Maci, Synthesis of modulated-metasurface antennas with amplitude, phase, and polarization control, *IEEE Trans. Antennas Propag.* **64**, 3907 (2016).
- [38] M. Faenzi, G. Minatti, D. González-Ovejero, F. Caminita, E. Martini, C. Della Giovampaola, and S. Maci, Metasurface antennas: New models, applications and realizations, *Sci. Rep.* **9**, 10178 (2019).
- [39] G. Minatti, F. Caminita, E. Martini, and S. Maci, Flat optics for leaky-waves on modulated metasurfaces: Adiabatic Floquet-wave analysis, *IEEE Trans. Antennas Propag.* **64**, 3896 (2016).
- [40] M. Bodehou, C. Craeye, E. Martini, and I. Huynen, A quasi-direct method for the surface impedance design of modulated metasurface antennas, *IEEE Trans. Antennas Propag.* **67**, 24 (2019).
- [41] S. N. Tsvetkova, D.-H. Kwon, A. Díaz-Rubio, and S. A. Tretyakov, Near-perfect conversion of a propagating plane wave into a surface wave using metasurfaces, *Phys. Rev. B* **97**, 115447 (2018).
- [42] E. Abdo-Sánchez, M. Chen, A. Epstein, and G. V. Eleftheriades, A leaky-wave antenna with controlled radiation using a bianisotropic Huygens' metasurface, *IEEE Trans. Antennas Propag.* **67**, 108 (2019).
- [43] M. Benzaouia, J. D. Joannopoulos, S. G. Johnson, and A. Karalis, Analytical Criteria for Designing Multiresonance Filters in Scattering Systems, with Application to Microwave Metasurfaces, *Phys. Rev. Applied* **17**, 034018 (2022).
- [44] F. Monticone and A. Alù, Leaky-wave theory, techniques, and applications: From microwaves to visible frequencies, *Proc. IEEE* **103**, 793 (2015).
- [45] Z. Lin, V. Liu, R. Pestourie, and S. G. Johnson, Topology optimization of freeform large-area metasurfaces, *Opt. Express* **27**, 15765 (2019).
- [46] Z. Lin, B. Groever, F. Capasso, A. W. Rodriguez, and M. Lončar, Topology-optimized multilayered metaoptics, *Phys. Rev. Applied* **9**, 044030 (2018).
- [47] J. Jiang and J. A. Fan, Simulator-based training of generative neural networks for the inverse design of metasurfaces, *Nanophotonics* **9**, 1059 (2020).
- [48] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, Inverse design in nanophotonics, *Nat. Photonics* **12**, 659 (2018).
- [49] H. Chung and O. D. Miller, High-na achromatic metalenses by inverse design, *Opt. Express* **28**, 6945 (2020).
- [50] E. Bayati, R. Pestourie, S. Colburn, Z. Lin, S. G. Johnson, and A. Majumdar, Inverse designed metalenses with extended depth of focus, *ACS Photonics* **7**, 873 (2020).
- [51] E. Bayati, R. Pestourie, S. Colburn, Z. Lin, S. G. Johnson, and A. Majumdar, Inverse designed extended depth of focus metaoptics for broadband imaging in the visible, *Nanophotonics* **11**, 2531 (2021).
- [52] Z. Li, R. Pestourie, J.-S. Park, Y.-W. Huang, S. G. Johnson, and F. Capasso, Inverse design enables large-scale high-performance meta-optics reshaping virtual reality, *Nat. Commun.* **13**, 2409 (2022).
- [53] T. W. Hughes, M. Minkov, V. Liu, Z. Yu, and S. Fan, Full wave simulation and optimization of large area metalens, in *OSA Optical Design and Fabrication 2021 (Flat Optics, Freeform, IODC, OFT)* (Optica Publishing Group, Washington, D.C., 2021), p. FTh3C.5.
- [54] T. W. Hughes, M. Minkov, V. Liu, Z. Yu, and S. Fan, A perspective on the pathway toward full wave simulation of large area metalenses, *Appl. Phys. Lett.* **119**, 150502 (2021).
- [55] J. Skarda, R. Trivedi, L. Su, D. Ahmad-Stein, H. Kwon, S. Han, S. Fan, and J. Vucković, Simulation of large-area metasurfaces with a distributed transition matrix method, [arXiv:2107.09879](https://arxiv.org/abs/2107.09879).
- [56] A. Mutapcic, S. Boyd, A. Farjadpour, S. G. Johnson, and Y. Avniel, Robust design of slow-light tapers in periodic waveguides, *Eng. Opt.* **41**, 365 (2009).
- [57] A. Oskooi, A. Mutapcic, S. Noda, J. D. Joannopoulos, S. P. Boyd, and S. G. Johnson, Robust optimization of adiabatic tapers for coupling to slow-light photonic-crystal waveguides, *Opt. Express* **20**, 21558 (2012).
- [58] R. Pestourie, Y. Mroueh, T. V. Nguyen, P. Das, and S. G. Johnson, Active learning of deep surrogates for PDEs: Application to metasurface design, *npj Comput. Mater.* **6**, 164 (2020).
- [59] M. L. Povinelli, S. G. Johnson, and J. D. Joannopoulos, Slow-light, band-edge waveguides for tunable time delays, *Opt. Express* **13**, 7145 (2005).
- [60] T. Felici and H. W. Engl, On shape optimization of optical waveguides using inverse problem techniques, *Inverse Probl.* **17**, 1141 (2001).
- [61] Z. Lin and S. G. Johnson, Overlapping domains for topology optimization of large-area metasurfaces, *Opt. Express* **27**, 32445 (2019).
- [62] T. Phan, D. Sell, E. W. Wang, S. Doshay, K. Edee, J. Yang, and J. A. Fan, High-efficiency, large-area, topology-optimized metasurfaces, *Light Sci. Appl.* **8**, 48 (2019).
- [63] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 2nd ed. (Cambridge University Press, Cambridge, 2017).
- [64] A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. Joannopoulos, and S. G. Johnson, Meep: A flexible

- free-software package for electromagnetic simulations by the FDTD method, *Comput. Phys. Commun.* **181**, 687 (2010).
- [65] S. G. Johnson and J. D. Joannopoulos, Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis, *Opt. Express* **8**, 173 (2001).
- [66] S. G. Johnson, M. Ibanescu, M. A. Skorobogatiy, O. Weisberg, J. D. Joannopoulos, and Y. Fink, Perturbation theory for Maxwell's equations with shifting material boundaries, *Phys. Rev. E* **65**, 066611 (2002).
- [67] L. D. Landau, L. P. Pitaevskii, and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Butterworth-Heinemann, Oxford, UK, 1984).
- [68] J. Skarda, R. Trivedi, L. Su, D. Ahmad-Stein, H. Kwon, S. Han, S. Fan, and J. Vučković, Low-overhead distribution strategy for simulation and optimization of large-area metasurfaces, *npj Comput. Mater.* **8**, 78 (2022).
- [69] Z. Gimbutas and L. Greengard, Fast multi-particle scattering: A hybrid solver for the Maxwell equations in microstructured materials, *J. Comput. Phys.* **232**, 22 (2013).
- [70] A. Egel, L. Pattelli, G. Mazzamuto, D. S. Wiersma, and U. Lemmer, Celes: Cuda-accelerated simulation of electromagnetic scattering by large ensembles of spheres, *J. Quantum Spectrosc. Radiat. Transfer* **199**, 103 (2017).
- [71] D. Pisssoort and F. Olyslager, Termination of periodic waveguides by PMLs in time-harmonic integral equation-like techniques, *IEEE Antennas Wireless Propag. Lett.* **2**, 281 (2003).
- [72] L. Zhang, J. H. Lee, A. Oskooi, A. Hochman, J. K. White, and S. G. Johnson, A novel boundary element method using surface conductive absorbers for full-wave analysis of 3-D nanophotonics, *J. Lightwave Technol.* **29**, 949 (2011).
- [73] O. P. Bruno and C. Pérez-Arancibia, Windowed Green function method for the Helmholtz equation in the presence of multiply layered media, *Proc. R. Soc. A* **473**, 20170161 (2017).