Efficient integral-equation-based method for accurate analysis of scattering from periodically arranged nanostructures

G. Kobidze, B. Shanker,* and D. P. Nyquist

2120 Engineering Building, Department of Electrical and Computer Engineering, Michigan State University,
East Lansing, Michigan 48824, USA

A plethora of applications are grounded on the physics of electromagnetic interaction with a periodic arrangement of nanostructures. These range from metamaterials and negative index materials to photonic band-gap structures to surface plasmon polariton optics to nanofrequency selective surfaces. There is therefore a need for rigorous physics based methods that are both accurate and fast to enable rapid design and analysis. Difficulties that need to be overcome to realize such a simulation tool are twofold: (i) at wavelengths in the range 200–1300 nm metals behave as dielectrics with negative real permittivity. Their frequency response must be explicitly accounted for in the simulation. (ii) The computational cost to compute response over a broad band of frequencies is high. This paper develops an integral-equation-based analysis technique that addresses these challenges. This integral equation relies on a periodic layered medium formulation. The Green's dyad for this formulation is derived, and separated into a superposition of two contributions: direct and reflected components. The means to accelerate the computation of the Green's dyad and the evaluation of inner products is prescribed. The proposed technique is validated extensively against available analytical data for hypothetical materials as well as silver. It is shown that this solver can accurately predict the enhanced transmission from perforated silver films for several configurations. While the application domain in this paper is the study of enhanced transmission in perforated silver films, the method presented herein is sufficiently general and can be applied to several other application domains with little or no change.

I. INTRODUCTION

Interest in analyzing scattering from periodically arranged scatterers has been on a steep rise, largely due to the perceived benefits of tailoring materials with specific electromagnetic properties [1,2]. For instance, research into metamaterials and negative index materials, photonic-band-gap structures, etc., has seen a meteoric rise in the past few years. In this paper, we intend developing techniques for efficiently analyzing electromagnetic response of periodically textured films. It will be assumed that the size of the perturbations is on the order of a few tenths to a hundred nanometers. The application domain that we will focus on is the study of the relation between surface plasmon polaritons (SPPs) and enhanced transmission. As will become apparent, the method prescribed herein can be seamlessly extended to other application domains.

SPPs are exponentially localized waves that exist at the interface between a dielectric and a metallic half space; it can be conveniently thought of as coupling between a surface wave and the electron oscillation of the metal. It was first demonstrated by Kretschman [3] and Otto [4] in the late 1960s, and later by several others [5–7], that these charge density oscillations exist at the interface between two materials whose real components of the permittivity have opposite signs. Examples wherein this may occur are at the interface between metal (semiconductor) and dielectric at optical frequencies. At optical wavelengths, the permittivity of metals is a function of the frequency of the incident electromagnetic wave, and the Drude free-electron model has been used to approximate the frequency dependance near the plasma frequency. When the frequency of the incident field is smaller than the plasma frequency, the real part of the relative permittivity can become less than −1, i.e., $\text{Re}\{\varepsilon_m\}<-1$. The interface between these two materials supports transverse magnetic (TM) plane wave surface modes. It can be shown [8] that the wave number of these modes is given by $k=k_0\sqrt{\varepsilon_m\varepsilon_d/(\varepsilon_m+\varepsilon_d)}$, where ε_m and ε_d are the relative permittivities of the metal and dielectric, respectively, and k_0 is the wave number in free space. For noble metals, the plasma frequency is in the optical regime; for instance, the plasma frequencies of silver and gold are 14×10^{15} s^{-1} and 13.8×10^{15} s^{-1}, respectively [9,10]. Gold is more lossy than silver. And, as a result, the propagation length of these modes at a gold-dielectric interface is considerably smaller. More rigorously, denoting the spectral wave numbers by k_x, k_y, and k_z and assuming that the metal-dielectric interface lies in the x-y plane, the existence of the SPP implies that k_x and k_y are large, and $k_z=\sqrt{k_0^2-k_x^2-k_y^2}$ will be large and imaginary. This condition restricts the wave to the interface; the propagation length is a function of the $\text{Im}(k_x)$ and $\text{Im}(k_y)$. In what follows, we assume that this is sufficiently smaller than $\text{Im}(k_z)$ so that one can sensibly discuss surface wave modes [11]. Indeed, this is not far from reality; it has been shown experimentally that propagation lengths in silver and gold at $\lambda=630$ nm are 19 and 3 μm, respectively. This figure increases substantially at a wavelength of 850 nm [12].

*Electronic address: bshanker@egr.msu.edu.; URL: http://www.egr.msu.edu/~bshanker

SPPs may be generated in a number of ways. They are always present to some extent simply due to thermal plasma fluctuations in the metal. Plasmons may also be generated directly by bombarding the metal with electrons. Some of those plasmons will couple to the vacuum electromagnetic field to form SPPs [6]. These modes may also be stimulated from the vacuum side by optically exciting the SPPs. Since the SPP wavelength is necessarily smaller than the vacuum wavelength of the exciting light, some means to generate superoscillatory waves must be used. These methods include prism coupling, high numerical aperture focusing through a high index coating, grating coupling [6], tapered probe coupling [13], and scattering from a small probe or surface defect [14]. In many cases the SPPs are generated at one side of a thin metal film by one of the aforementioned means and are subsequently observed on the other side of the film. The change in the geometry from the half space to the thin film affects little of the physics except that it results in lower losses to the bulk metal and thus larger propagation lengths.

Thus far, we have largely been discussing SPPs on planar interfaces. Exciting a surface wave on planar surfaces is possible only if the angle of incidence is close to the critical angle. One of the first methods proposed to launch a SPP was based on this fact, however, surface waves are also generated whenever the surface is textured in a periodic fashion. While the former can be analyzed using simple analytical models, the latter is not easily amenable to analysis using simplified models. This has not stopped assumptions that simplify analysis: For instance, given that SPPs are localized and are due to one polarization, they may be modeled using the two-dimensional (2D) scalar Helmholtz equation (infinite in y, another assumption) [15]. While scalar models are limited in their ability to model defects which result in additional polarization states and out-of-plane modes, they do yield some insight in the behavior of SPPs. Other approximations such as using an impedance boundary condition [16] and optimization of coupling to out-of-plane modes [17] have been published as well.

A range of techniques that have been developed for the analysis of gratings may also be used for SPP analysis. However, it is important to realize each method's advantages and, more importantly, its limitations. This said, the methods that have been used fall into two broad categories: (i) differential methods (and methods that have evolved from this technique), and (ii) full-wave methods. In what follows, we shall briefly overview both sets of techniques, and point out their limitations.

The classical differential method is essentially a spectral technique that relies on using the Fourier basis to represent each component of the field [18]. The coefficient for this basis is obtained by imposing boundary conditions. The differential method may be a viable tool for addressing many if not all types of gratings [19–21]. The modal method, that relies on solving Maxwell's equation in each rectangular domain in closed form and linking all domains via boundary conditions, imposes severe restrictions [19], and is most effective for gratings that are easily represented in terms of rectangular domains. Another evolution of this technique is the coupled wave method also called the Fourier modal method [22–24]. This technique relies on dividing the domain into rectangular slices and expressing solution as a sum of modes and then solving for these coefficients by imposing boundary conditions. As in the classical differential method, this method works very well when few Fourier harmonics are used. Finally, the coordinate transform method (or C method) introduces a different coordinate system that transforms corrugated surfaces into planar surfaces. This transformation not only simplifies the application of boundary conditions but also transforms Maxwell's equations in Fourier space into a matrix eigenvalue problem [25]. This method has been shown to work well for deep gratings as well as anisotropic ones [26]. As is evident from the above discussion, the roots of the classical differential equation methods are founded in the spectral method. While these methods are certainly powerful for a range of problems, one must be aware of their limitations as well. These are well known, and are repeated here for completion purposes only: They are as follows: (i) the number of terms required for convergence is dependent on the spatial support of discontinuity; (ii) staircasing approximation for profiles that are not rectangular; (iii) poor convergence for TM polarizations. Intense research has focused on overcoming all of these limitations. Indeed, the number of terms required for convergence can be overwhelming for two-dimensional periodicities.

The other broad class of methods that exist are the so-called full-wave methods. These methods involve disretization of fields in terms of local basis functions. Methods such as finite difference, finite elements, and integral equation based techniques fall under this category. Purely full-wave models of SPPs are scant. Most of these models use finite difference time domain (FDTD), which is ideal for proof of concept studies as it is well known that FDTD suffers from severe dispersion errors when the electrical size of the object is large. Note that while the object is small compared to free-space wavelength, modeling the field inside the metal requires a large number of unknowns; this is especially true near resonance. Also, it is well known that unless special techniques are used, FDTD is unstable for oblique angles of incidence [27], and methods prescribed in Ref. [28] are only valid at discrete angles of incidence. Integral equation based techniques offer considerably more accurate models albeit at a higher computational cost. At radio frequency (RF) frequencies, the extant literature on integral equation based analysis of periodic structures and the range of applications is vast; see the excellent book by Munk [29]. Likewise, it has been shown that integral equation based approaches may indeed be the method of choice if one is to make accurate predictions for arbitrarily shaped grating profiles [30–35]. The solution to these integral equations may be obtained in one of two ways: (i) by discretizing the equations using entire domain basis functions. While this is a possibility, the spectral summations are slowly converging especially when the observation and the source are in the same horizontal plane; (ii) alternatively, we can use basis functions that have finite spatial support, and convolve them with the periodic Green's function. The latter technique enables the development of acceleration procedures for the convergence of the spectral summation. The use of a volume integral equation using spatial basis functions for analyzing SPPs in nanostructures is not new; it has been used to analyze SPPs at
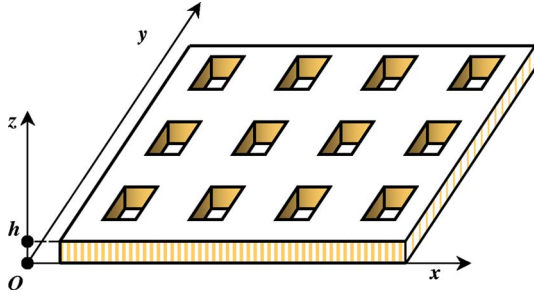
FIG. 1. (Color online) Thin noble metal film perforated periodically in the x and y directions.

interfaces and in tips of probes, etc., by Martin's group [36–39]. Our efforts to analyze SPPs in periodic structures has been restricted to using surface integral equations [40]. Both surface and volume integral methods suffer from a very high computational overhead. This is due to the fact that (i) the spatial variation of the fields in the dielectric region must be captured to high accuracy. It implies that when using a space of linear functions one needs at least ten samples per wavelength in the dielectric; (ii) the region to be represented is the dielectric region that can be electrically large.

Consequently, the principal contribution is twofold: (i) an alternative formulation is proposed that will considerably reduce the number of unknowns; (ii) a procedure is prescribed for accurate evaluation of the inner products. Two pleasant side effects of the proposed technique are that it enables rapid evaluation of all inner products for each spectral mode and it results in a more rapidly converging spectral sum. The methods developed here are not specific to the analysis of SPPs. They can be applied with no change to analyze other applications that exploit periodicity.

The layout of this paper is as follows: Section II will formulate the problem and detail the numerical implementation of the proposed approach. Discrete implementation of these integral equations will be detailed in Sec. III and techniques (both analytical and numerical) to rapidly evaluate these will be presented in Sec. IV. Section V will present a plethora of results that will serve to validate the proposed approach as well as demonstrate its accuracy and ability to predict the existence of enhanced transmission in periodically perforated structures. In Sec. VI, we will summarize the contribution of this paper and chart out directions for future research. Finally, we will formally complete derivation of the periodic layered medium Green's function in Appendix A, and prescribe the formulas necessary to analytically evaluate the integral over the Green's function in Appendix B.

II. FORMULATION

A. Preliminaries

The geometry being analyzed is described as follows: an infinite film whose bottom surface is flush with the $z=0$ plane has periodical perturbations, such as perforated holes, and is illuminated by a plane wave; see Fig. 1. The objective is to find the reflected and transmitted field. The perforation can be of any shape. We will assume that the thickness of the film is h, it occupies a region $0 \leq z \leq h$, and its permittivity and permeability are denoted by $\{\varepsilon_p, \mu_0\}$. Without loss of generality, it is assumed that free space exists above this configuration. The volume occupied by all the perforations is denoted by $\Omega = \cup_{mn} \Omega_{mn}$. It is obvious that Ω is a multiply connected domain. Again, without loss of generality, we will assume that the mother cell is periodically replicated along the x and y axes with a period D_x and D_y, respectively. To solve for the scattered fields, we adopt a two-step process: (i) using the volume equivalence theorem, the problem can be reformulated as follows: Find the fields radiated by equivalent currents $\mathbf{J}(\mathbf{r}) = -i\omega(\varepsilon_0 - \varepsilon_p)\mathbf{E}(\mathbf{r}) \forall \ \mathbf{r} \in \Omega$ in the presence of multilayered medium; (ii) we will exploit the fact that the configuration is periodic to reduce the computational domain to the mother cell. It should be noted that while we have used the volume equivalence theorem, we could have used the surface equivalence theorem to formulate an equivalent problem; however, other difficulties emerge that will be elaborated upon later in this paper.

Denoting the incident and scattered electric fields by $\mathbf{E}^{inc}(\mathbf{r})$ and $\mathbf{E}^{sca}(\mathbf{r})$, the total electric field may be written $\forall \ \mathbf{r} \in \Omega$ as

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + \mathbf{E}^{sca}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + \mathcal{G}(\mathbf{r}) \star \mathbf{J}(\mathbf{r}), \quad (1)$$

where $\mathcal{G}(\mathbf{r})$ is the dyadic Green's function, and $\star$ denotes a spatial convolution *and* a dot product. In writing Eq. (1), we have been lax in our notation (largely due to convenience) as $\mathcal{G}(\mathbf{r})$ has a source point singularity. Methods of overcoming this are well known, and will be referred to later in the text. Note that the incident field is defined as the field that exists in a multilayered medium in the absence of perforations. Assume that the origin is located at a corner of the *mother cell* Ω_{00}. If the upgoing field in the first layer is of the form $\mathbf{E}^{up}(\mathbf{r}) = \mathbf{E}_0 \exp[i\mathbf{k}^{inc} \cdot \mathbf{r}]$, it is apparent that the currents in cell Ω_{00} and those in Ω_{mn} are related by $\mathbf{J}_{mn}(\mathbf{r}) = \mathbf{J}_{00 - \boldsymbol{\rho}_{mn}}(\mathbf{r}) \exp[i\mathbf{k}^{inc} \cdot \boldsymbol{\rho}_{mn}]$, where $\boldsymbol{\rho}_{mn} = mD_x \hat{x} + nD_y \hat{y}$. It is immediately apparent that the current can be written as

$$\mathbf{J}(\mathbf{r}) = \mathbf{J}_{00}(\mathbf{r}) \otimes P(\mathbf{r}),$$

$$P(\mathbf{r}) = \sum_m \sum_n \delta(\mathbf{r} - \boldsymbol{\rho}_{mn}) \exp[i\mathbf{k}^{inc} \cdot \mathbf{r}], \quad (2)$$

where $\otimes$ denotes a convolution, supp $[\mathbf{J}_{00}(\mathbf{r})] \in \bar{\Omega}_{00}$, and $\bar{\Omega}_{00}$ denotes the completion of Ω_{00}. Using the above expression in Eq. (1) and the Poisson summation formula, it can be shown that

$$\mathcal{G}(\mathbf{r}) \otimes \mathbf{J}(\mathbf{r}) = \mathcal{G}(\mathbf{r}) \otimes \mathbf{J}_{00}(\mathbf{r}) \otimes P(\mathbf{r}) \quad (3a)$$

$$= \mathbb{F}^{-1}\{\mathbb{F}(\mathcal{G}) \times \mathbb{F}(P(\mathbf{r}))\} \otimes \mathbf{J}_{00}(\mathbf{r}) \quad (3b)$$

$$= \frac{1}{D_x D_y} \left[\sum_m \sum_n \mathbb{G}(k_{xm}, k_{yn}, z) \exp[i\mathbf{k}_{mn} \cdot \mathbf{r}] \right]$$

$$\otimes \mathbf{J}_{00}(\mathbf{r}) \tag{3c}$$

$$= \widetilde{\mathcal{G}}(\mathbf{r}) \otimes \mathbf{J}_{00}(\mathbf{r}). \tag{3d}$$

In the above equations, $\mathbb{F}$ denotes the Fourier transform (and its inverse), $\mathbb{G}(k_{xm}, k_{ym}, z)$ denotes a Fourier transform of the dyadic Green's function, $\widetilde{\mathcal{G}}(\mathbf{r})$ is the periodic Green's function, $k_{xm} = 2\pi m/D_x + k_x^i$, $k_{yn} = 2\pi n/D_y + k_y^i$, $k_{zmn} = \sqrt{k^2 - k_{xm}^2 - k_{yn}^2}$, $k_x^i = \mathbf{k}^{inc} \cdot \hat{x}$, $k_y^i = \mathbf{k}^{inc} \cdot \hat{y}$ and $\mathbf{k}_{mn} = k_{xm}\hat{x} + k_{yn}\hat{y} + k_{zmn}\hat{z}$. Henceforth a *tilde* over a quantity will denote the periodic version of the same.

Thus, in the above, we have expressed the periodic Green's function as a summation over the spectral components of layered medium Green's function. This is fortuitous as the starting point for the derivation of the Green's dyad for a source embedded in a layered medium is in spectral form—thus expressions required for deriving the Green's functions are readily available.

B. Layered medium Green's function

Research into the construction of layered medium Green's functions has been extensive. We refer the reader to an excellent survey of literature in this field by Michalski [41]. As was pointed out in the last section, the derivation of the periodic layered medium Green's function hinges on the expressions used for deriving the layered medium Green's function. For completeness, we will very briefly review the derivation of the layered medium Green's function. Our derivation will follow that of Chew [42] and Kong [43] largely due to familiarity. Observe that the unbounded scalar Green's function can be written in the spectral form using the Weyl identity

$$g(R) = \frac{e^{ikR}}{4\pi R} = \frac{i}{8\pi^2} \int \int_{-\infty}^{\infty} dk_x dk_y \frac{e^{ik_x(x-x')+ik_y(y-y')+ik_z|z-z'|}}{k_z}, \tag{4}$$

where $k^2 = k_0^2(\varepsilon_r \mu_r)$, $k_z = \text{sgn}(k_z)\sqrt{k^2 - k_\rho^2}$, $k_\rho^2 = k_x^2 + k_y^2$, and $\text{sgn}(k_z) \in \{-1, 1\}$ is chosen such that $\text{Im}(k_z) \geqslant 0$. This expression has been the starting point for analyzing fields due to sources above a layered media using either Hertzian potentials or TM/TE field decomposition [42–44]. Alternatively, one can derive boundary conditions on the vector Hertz potentials as well [45–47]. The latter avoids the source point singularity inherent in the other two formulations. Here, we shall use the decomposition of fields into transverse electric (TE) and TM modes to derive the Green's function, and then separate the singular principal part that corresponds to the head wave, and the nonsingular reflected part that corresponds to the lateral wave.

For completeness, the derivation and integral expressions for the total scattered, reflected, and transmitted fields in the layered medium are given in Appendix A. Assuming that the periodic layered medium Green's function can be written as

a superposition of the direct and reflected Green's functions, $\widetilde{\mathcal{G}}(\mathbf{r}) = \widetilde{\mathcal{G}}^d(\mathbf{r}) + \widetilde{\mathcal{G}}^r(\mathbf{r})$, the total scattered electric field in the *periodic* layered medium can then be expressed as

$$\mathbf{E}^{sca}(\mathbf{r}) = \mathbf{E}^d(\mathbf{r}) + \mathbf{E}^r(\mathbf{r}) = \widetilde{\mathcal{G}}^d(\mathbf{r}) \star \mathbf{J}_{00}(\mathbf{r}) + \widetilde{\mathcal{G}}^r(\mathbf{r}) \star \mathbf{J}_{00}(\mathbf{r}) \tag{5}$$

The direct Green's function $\widetilde{\mathcal{G}}^d(\mathbf{r}) = i\omega\mu[\mathcal{I} + (1/k_p^2)\boldsymbol{\nabla}\boldsymbol{\nabla}]\tilde{g}^d(\mathbf{r})$, where $\mathcal{I}$ is the idempotent, k_p^2 is the wave number in the medium, and $\tilde{g}^d(\mathbf{r})$ is the periodic scalar Green's function that satisfies Sommerfeld radiation boundary condition with wave number k_p. Parenthetically, we note that the the expression in Eq. (5) for the evaluation of the direct field cannot be computed numerically as it has a source point singularity. Thus its use here is purely for notational convenience as we shall elaborate the techniques of evaluating the direct field using a mixed potential formulation. Denoting $k_{\rho mn}^2 = k_{xm}^2 + k_{yn}^2$ and $e_{mn} = (1/8\pi^2 D_x D_y k_{zpmn}\omega\varepsilon_p)e^{ik_{xm}(x-x')+ik_{yn}(y-y')}$, the components of the reflected Green's dyadic $\widetilde{\mathcal{G}}^r(\mathbf{r})$ can be expressed as

$$\widetilde{\mathcal{G}}_{xx}^r = -\sum_{m=-M_F}^{M_F} \sum_{n=-N_F}^{N_F} e_{mn} \frac{k_{pzmn}^2 k_{xm}^2 F_r^{TM}(z,z') + k_{yn}^2 k_p^2 F_r^{TE}(z,z')}{k_{\rho mn}^2}, \tag{6a}$$

$$\widetilde{\mathcal{G}}_{xy}^r = -\sum_{m=-M_F}^{M_F} \sum_{n=-N_F}^{N_F} e_{mn} k_{xm} k_{yn} \frac{k_{pzmn}^2 F_r^{TM}(z,z') - k_p^2 F_r^{TE}(z,z')}{k_{\rho mn}^2}, \tag{6b}$$

$$\widetilde{\mathcal{G}}_{xz}^r = -\sum_{m=-M_F}^{M_F} \sum_{n=-N_F}^{N_F} e_{mn} k_{xm} k_{pzmn} dF_r^{TM}(z,z'), \tag{6c}$$

$$\widetilde{\mathcal{G}}_{yx}^r = \widetilde{\mathcal{G}}_{xy}^r, \tag{6d}$$

$$\widetilde{\mathcal{G}}_{yy}^r = -\sum_{m=-M_F}^{M_F} \sum_{n=-N_F}^{N_F} e_{mn} \frac{k_{pzmn}^2 k_{yn}^2 F_r^{TM}(z,z') + k_{xm}^2 k_p^2 F_r^{TE}(z,z')}{k_{\rho mn}^2}, \tag{6e}$$

$$\widetilde{\mathcal{G}}_{yz}^r = -\sum_{m=-M_F}^{M_F} \sum_{n=-N_F}^{N_F} e_{mn} k_{yn} k_{pzmn} dF_r^{TM}(z,z'), \tag{6f}$$

$$\widetilde{\mathcal{G}}_{zx}^r = -\widetilde{\mathcal{G}}_{xz}^r, \tag{6g}$$

$$\widetilde{\mathcal{G}}_{zy}^r = -\widetilde{\mathcal{G}}_{yz}^r, \tag{6h}$$

$$\widetilde{\mathcal{G}}_{zz}^r = -\sum_{m=-M_F}^{M_F} \sum_{n=-N_F}^{N_F} e_{mn} k_{\rho mn}^2 F_r^{TM}(z,z'). \tag{6i}$$

In the above, the quantities F_r^{TM}, F_r^{TE}, and dF_r are defined in Appendix A in terms of the generalized TE and TM reflection coefficients, and the integers M_F and N_F correspond to the number of terms that are necessary to take into account to

achieve the desired level of accuracy in computing the spectral sum.

III. DISCRETE SOLUTION OF THE INTEGRAL EQUATION

To solve the problem numerically, one needs to represent the unknown equivalent current density $\mathbf{J}(\mathbf{r}) = -i\omega\kappa(\mathbf{r})\mathbf{D}(\mathbf{r})$ in terms of spatial basis functions. Here $\kappa(\mathbf{r}) = [\varepsilon_0 - \varepsilon_p(\mathbf{r})/\varepsilon_0]$ is the contrast ratio. Following the usual practice [48,49], we represent the flux density in terms of spatial basis functions such that

$$\mathbf{D}_{00}(\mathbf{r}) = \sum_{j=1}^{N} D_j \mathbf{f}_j(\mathbf{r}). \tag{7}$$

The basis functions $\mathbf{f}(\mathbf{r})$ are divergence conforming, i.e., they have a finite divergence. More specifically, the basis functions are defined on a set of tetrahedra. Each basis function $\mathbf{f}_j(\mathbf{r})$ is associated with one facet S_j that can be shared by at most two tetrahedra V_j^+ and V_j^- and these (at most two) adjacent tetrahedra constitute the domain of support for $\mathbf{f}_j(\mathbf{r})$, i.e., $\mathrm{supp}\{\mathbf{f}_j(\mathbf{r})\} = V_j = V_j^+ \cup V_j^-$. The basis functions are defined as $\mathbf{f}_j(\mathbf{r}) = \mathbf{f}_j^{\pm}(\mathbf{r})$, with $\mathbf{f}_j^{\pm} = \pm (a_j/3v_j^{\pm})(\mathbf{r} - \mathbf{r}_j^{\pm}) \forall~\mathbf{r} \in V_j^{\pm}$. In the above expressions, a_j is the area of S_j, $v_j^{\pm}$ is the volume of $V_j^{\pm}$, and $\mathbf{r}_j^{\pm}$ is the vertex of $V_j^{\pm}$ opposite the facet. Also, the permittivity is assumed to be piecewise constant. Using the above defined representation of the flux density in Eqs. (1) and (5) and using Galerkin testing results in

$$\langle \mathbf{E}, \mathbf{f}_i \rangle = \langle \mathbf{E}^{inc}, \mathbf{f}_i \rangle + \langle \mathbf{E}^d, \mathbf{f}_i \rangle + \langle \mathbf{E}^r, \mathbf{f}_i \rangle, \tag{8}$$

$\forall~i = 1, \ldots, N$ results in a system of matrix equations. To evaluate the direct field $\mathbf{E}^d(\mathbf{r})$, we espouse the mixed potential approach. Details of this approach have been expounded extensively in Ref. [48], and has successfully been applied to both time and frequency domain simulation of scattering from inhomogeneous, dispersive or anisotropic targets [49,50], albeit not explicitly using the periodic Green's function (free-space periodic Green's function has been implicitly used in Ref. [49] to validate the presented methodology). More specifically,

$$\mathbf{E}^d(\mathbf{r}) = i\omega\mathbf{A}^d(\mathbf{r}) - \mathbf{\nabla}\Phi^d(\mathbf{r}), \tag{9a}$$

$$\mathbf{A}^d(\mathbf{r}) = \mu \widetilde{g}^d(\mathbf{r}) \star \mathbf{J}_{00}(\mathbf{r}), \tag{9b}$$

$$\Phi^d(\mathbf{r}) = \frac{1}{\varepsilon_p} \widetilde{g}^d(\mathbf{r}) \star \sigma_{00}(\mathbf{r}). \tag{9c}$$

Here, $\sigma_{00}(\mathbf{r})$ denotes the charge density and comprises both volumetric and surface charge densities that are related to the current as $(i/\omega)\mathbf{\nabla} \cdot \mathbf{J}_{00}(\mathbf{r})$ and $-(i/\omega)\hat{n} \cdot \mathbf{J}_{00}(\mathbf{r})$, respectively. Also, $\hat{n}$ denotes the outward pointing normal to the volume Ω_{00}. While this representation does indeed reduce the order of singularity associated with the scalar potential, we note that it can be further reduced using the fact that $\langle \mathbf{\nabla}\Phi(\mathbf{r}), \mathbf{f}_j \rangle = \oint_{\partial V_j} d\mathbf{r}~\Phi(\mathbf{r})\hat{n} \cdot \mathbf{f}_j(\mathbf{r}) - \langle \Phi(\mathbf{r})\mathbf{\nabla} \cdot \mathbf{f}_i \rangle$, where ∂V_j denotes the surface bounding V_j. With these manipulations, it is possible

to evaluate the inner products. Means of evaluating the integral analytically are given in Ref. [48] for the scalar Green's function. Note that none of these integrals are straightforward and special techniques must be developed to evaluate these rapidly and accurately. These are elaborated in the next subsection, as is the need to develop exact integration techniques for evaluating $\langle \mathbf{E}^r, \mathbf{f}_i \rangle$. Thus Galerkin testing results in a matrix equation of the form

$$\mathcal{Z}\,\mathcal{I} = \mathcal{F}, \tag{10}$$

where $\mathcal{I}_j = D_j$, $\mathcal{F}_j = \int_{V_j} d\mathbf{r}~\mathbf{E}^{inc}(\mathbf{r}) \cdot \mathbf{f}_j$, and

$$
\begin{aligned}
\mathcal{Z}_{ij} &= \int_{V_i} d\mathbf{r}\, \frac{\mathbf{f}_i(\mathbf{r}) \cdot \mathbf{f}_j(\mathbf{r}')}{\epsilon_0} \\
&+ i\omega\kappa_j \int_{V_i} d\mathbf{r}~\mathbf{f}_i(\mathbf{r}) \cdot \int_{V_j} d\mathbf{r}'\, \widetilde{\mathcal{G}}^r(\mathbf{r},\mathbf{r}') \cdot \mathbf{f}_j(\mathbf{r}') \\
&- \mu_0\omega^2\kappa_j \int_{V_i} d\mathbf{r}~\mathbf{f}_i(\mathbf{r}) \cdot \int_{V_j} d\mathbf{r}'\mathbf{f}_j(\mathbf{r}')\widetilde{g}(R) \\
&+ \frac{\kappa_j}{\varepsilon_p} \int_{V_i} d\mathbf{r}~\mathbf{\nabla} \cdot \mathbf{f}_i(\mathbf{r}) \int_{V_j} d\mathbf{r}'\mathbf{\nabla}' \cdot \mathbf{f}_j(\mathbf{r}')\widetilde{g}(R) \\
&- \frac{\kappa_j}{\varepsilon_p} \int_{V_i} d\mathbf{r}~\mathbf{\nabla} \cdot \mathbf{f}_i(\mathbf{r}) \oint_{S_j} d\mathbf{r}'\mathbf{f}_j(\mathbf{r}') \cdot \hat{n}'\widetilde{g}(R) \\
&- \frac{\kappa_j}{\varepsilon_p} \oint_{S_i} d\mathbf{r}~\mathbf{f}_i(\mathbf{r}) \cdot \hat{n} \int_{V_j} d\mathbf{r}'\mathbf{\nabla}' \cdot \mathbf{f}_j(\mathbf{r}')\widetilde{g}(R) \\
&+ \frac{\kappa_j}{\varepsilon_p} \oint_{S_i} d\mathbf{r}~\mathbf{f}_i(\mathbf{r}) \cdot \hat{n} \oint_{S_j} d\mathbf{r}'\mathbf{f}_j(\mathbf{r}') \cdot \hat{n}'\widetilde{g}(R).
\end{aligned} \tag{11}
$$

In these equations, $R = \mathbf{R} = |\mathbf{r} - \mathbf{r}'|$ and $S = \partial V$ denotes the surface bounding the volume V. The accurate and rapid evaluation of these integrals is critical efficient analysis of periodic structure. The first integral in this equation is readily evaluated analytically as it is simply an inner product of two basis functions [48,49]. Different methods will be used for the two types of Green's functions encountered in these integrals, and they are elucidated next.

IV. RAPID AND EFFICIENT EVALUATION OF INNER PRODUCTS

Two types of periodic Green's functions are encountered in Eq. (11)—the scalar Green's function for homogeneous media and the reflected Green's dyadic. They need to be treated differently when evaluating inner products. In what follows, we will analyze these in sequence.

A. Evaluation of scalar periodic Green's function

The scalar periodic Green's function contains a source point singularity that must be accounted for rigorously. As was mentioned earlier, this is done using well established techniques [51]. Typically, one uses a numerical quadrature to evaluate the integrals in Eq. (11) This implies that the periodic Green's function needs to be rapidly evaluated.

Three methods exist for evaluating the Green's function; (i) direct summation, (ii) spectral summation, and (iii) Ewald summation. As is well known [52], the first two converge slowly at approximately the same rate when $z - z' = 0$. The only exception to this rule is when the material of the layer is lossy, in which case the direct sum converges exponentially. For finite $z - z'$, the convergence of the spectral sum is exponential as well. When this is not the case, the convergence is prohibitively slow, and one typically uses the Ewald transformation which makes the sum exponentially convergent. The Ewald transform uses the definition of the error function $\text{erf}(u) = (2/\sqrt{\pi})\int_0^u e^{-t^2} dt$ and the complementary error function $\text{erfc}(u) = 1 - \text{erf}(u) = (2/\sqrt{\pi})\int_u^\infty e^{-t^2} dt$, as well as the identity

$$\frac{e^{ikR}}{R} = \frac{2}{\sqrt{\pi}}\left(\int_0^E + \int_E^\infty\right) e^{-R^2 t^2 + (k^2/4t^2)} dt, \quad (12)$$

where the parameter E is chosen appropriately to balance the computational cost between the two integrals. In this case $\tilde{g}(\mathbf{r}, \mathbf{r}') = \tilde{g}_1(\mathbf{r}, \mathbf{r}') + \tilde{g}_2(\mathbf{r}, \mathbf{r}')$, where

$$\tilde{g}_1 = -\frac{1}{D_x D_y} \sum_{m=-\infty}^\infty \sum_{n=-\infty}^\infty \left[e^{\gamma R_z}\text{erfc}\left(\frac{-\gamma}{2E} - R_z E\right) \right.$$
$$\left. + e^{-\gamma R_z}\text{erfc}\left(\frac{-\gamma}{2E} + R_z E\right)\right]\frac{e^{ik_x R_x + ik_y R_y}}{4\gamma} \quad (13)$$

and

$$\tilde{g}_2 = \sum_{m=-\infty}^\infty \sum_{n=-\infty}^\infty \left[e^{-ikR_{mn}}\text{erfc}\left(R_{mn}E - \frac{ik}{2E}\right) \right.$$
$$\left. + e^{ikR_{mn}}\text{erfc}\left(R_{mn}E + \frac{ik}{2E}\right)\right]\frac{e^{imk_x^{inc}D_x + ink_y^{inc}D_y}}{8\pi R_{mn}}, \quad (14)$$

where R_i for $i \in \{x, y, z\}$ denotes a component of $\mathbf{R} = \mathbf{r} - \mathbf{r}'$, and $k_z = \text{Re}\{k_z\} + i\text{Im}\{k_z\} = \beta_z + i\alpha_z$,

$$\gamma = ik_z = \begin{cases} i\beta_z & \text{if} \quad k_z^2 \geq 0 \\ -\alpha_z & \text{if} \quad k_z^2 < 0 \end{cases}. \quad (15)$$

Note that instead of computing two complex-valued complementary error functions, it is considerably easier to compute one real component $K(\zeta)$ or one imaginary component $L(\zeta)$ of one Faddeeva function $w(\zeta) = w(\zeta_{Re} + i\zeta_{Im}) = e^{-\zeta^2}\text{erfc}(-i\zeta) = K(\zeta) + iL(\zeta)$, using very efficient Humliček [53–55] algorithms. Let $\zeta = (\beta_z/2E) - iR_z E$. Then $(-i\beta_z/2E) - R_z E = -i\zeta$, $\zeta^2 = (\beta_z^2/4E^2) - R_z^2 E^2 - i\beta_z R_z$, and

$$e^{i\beta_z R_z}\text{erfc}\left(\frac{-i\beta_z}{2E} - R_z E\right) = w(\zeta)e^{(\beta_z^2/4E^2 - r_z^2 E^2)},$$

$$e^{-i\beta_z R_z}\text{erfc}\left(\frac{-i\beta_z}{2E} + R_z E\right) = w(\bar{\zeta})e^{(\beta_z^2/4E^2 - R_z^2 E^2)}. \quad (16)$$

Using the symmetry property of the Faddeeva function $w(\bar{\zeta}) = 2e^{-\zeta^2} - w(\zeta)$, the sum of two functions with complex-conjugate arguments is written as $w(\zeta) + w(\bar{\zeta}) = 2[iL(\zeta) + e^{-\zeta^2}]$, such that

$$e^{i\beta_z R_z}\text{erfc}\left(\frac{-i\beta_z}{2E} - R_z E\right) + e^{-i\beta_z R_z}\text{erfc}\left(\frac{-i\beta_z}{2E} + R_z E\right)$$
$$= 2\left[ie^{(\beta_z^2/4E^2 - R_z^2 E^2)}L(\zeta) + e^{-i\beta_z R_z}\right]. \quad (17)$$

Then the first term of the Green's function becomes for $k_z^2 \geq 0$

$$\tilde{g}_1 = \frac{1}{D_x D_y} \sum_{m=-\infty}^\infty \sum_{n=-\infty}^\infty \left[ie^{-i\beta_z R_z} + e^{(\beta_z^2/4E^2 - R_z^2 E^2)}L \right.$$
$$\left. \times \left(\frac{\beta_z}{2E} - iR_z E\right)\right]\frac{e^{ik_x R_x + ik_y R_y}}{2\beta_z} \quad (18)$$

and for $k_z^2 < 0$

$$\tilde{g}_1 = \frac{1}{D_x D_y} \sum_{m=-\infty}^\infty \sum_{n=-\infty}^\infty \left[e^{\alpha_z R_z}\text{erfc}\left(\frac{\alpha_z}{2E} + R_z E\right) \right.$$
$$\left. + e^{-\alpha_z R_z}\text{erfc}\left(\frac{\alpha_z}{2E} - R_z E\right)\right]\frac{e^{ik_x R_x + ik_y R_y}}{4\alpha_z}. \quad (19)$$

Similar manipulations can be used to evaluate the second term. Using $\zeta = k/2E + iR_{mn}E$, $R_{mn}E - ik/2E = -i\zeta$, $\zeta^2 = k^2/4E^2 - R_{mn}^2 E^2 + ikR_{mn}$, results in

$$e^{-ikR_{mn}}\text{erfc}\left(R_{mn}E - \frac{ik}{2E}\right) = w(\zeta)e^{(k^2/4E^2 - R_{mn}^2 E^2)},$$

$$e^{ikR_{mn}}\text{erfc}\left(R_{mn}E + \frac{ik}{2E}\right) = w(-\bar{\zeta})e^{(k^2/4E^2 - R_{mn}^2 E^2)}. \quad (20)$$

Another symmetry property of the Faddeva function $w(-\bar{\zeta}) = \overline{w(\zeta)}$ can be used to express the sum of two functions with negative complex-conjugate arguments as $w(\zeta) + w(-\bar{\zeta}) = 2K(\zeta)$ with

$$e^{-ikR_{mn}}\text{erfc}\left(R_{mn}E - \frac{ik}{2E}\right) + e^{ikR_{mn}}\text{erfc}\left(R_{mn}E + \frac{ik}{2E}\right)$$
$$= 2K\left(\frac{k}{2E} + iR_{mn}E\right)e^{(k2/4E2 - R2_{mn}E2)} \quad (21)$$

and the second term becomes

$$\tilde{g}_2 = \sum_{m=-\infty}^\infty \sum_{n=-\infty}^\infty e^{(k^2/4E^2 - R_{mn}^2 E^2)}K\left(\frac{k}{2E} + iR_{mn}E\right)\frac{e^{imk_x^{inc}D_x + ink_y^{inc}D_y}}{4\pi R_{mn}}. \quad (22)$$

Using these Faddeeva functions instead of error functions in the Ewald series considerably accelerates the computation of the Green's function. For instance, a 20-fold speedup was obtained for computing the Green's function within a relative error of 10^{-4}.

B. Evaluation of inner products with the reflected Green's dyad

Next, we will elaborate on the evaluation of the inner product with the reflected Green's dyad in Eq. (11). The

methods presented here can also can be used to evaluate the other inner products in Eq. (11) whenever the Green's function is expressed in the spectral form. Indeed, it is highly recommended that the procedure prescribed in this section be used as it is considerably more accurate; see Sec. V for convergence data.

1. Local integral representation

The accurate evaluation of

$$\mathcal{Z}_{ij}^{r} = \int_{V_i} d\mathbf{r}\, \mathbf{f}_i(\mathbf{r}) \cdot \int_{V_j} d\mathbf{r}'\, i\omega \widetilde{\mathcal{G}}^{r}(\mathbf{r},\mathbf{r}') \cdot \mathbf{f}_j(\mathbf{r}') \qquad (23)$$

using numerical quadrature rules does not appear feasible as integrands that correspond to higher spectra oscillate in such a manner that they cannot be accurately described by polynomials or be reconstructed accurately without satisfying the Nyquist criterion. Hence analytical or close-form integration is necessary.

Any component of Eq. (6) can be expressed in the form

$$\widetilde{\mathcal{G}}_{\xi,\eta}^{r}(\mathbf{r},\mathbf{r}') = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} G_{\xi,\eta}^{mn}, \qquad (24)$$

where $G_{\xi,\eta}^{mn}$ are the integrands from Eq. (6) ξ and η being any from the set of indices $\{x,y,z\}$, and the superscript mn maps to the spectral components with $k_x = k_{xm}$, $k_y = k_{yn}$, and $k_{zp} = k_{pznm}$. A close look at Eq. (6) reveals that for every spectral component, each element of the reflected part of the Green's function can be represented as a sum of four terms,

$$\omega G_{\xi\eta}^{mn} = \mathcal{A}_{\xi\eta}^{mn} e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(z+z')]}$$

$$+ \mathcal{B}_{\xi\eta}^{mn} e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(2h+z-z')]}$$

$$+ \mathcal{C}_{\xi\eta}^{mn} e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(2h-z+z')]}$$

$$+ \mathcal{D}_{\xi\eta}^{mn} e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(2h-z-z')]}, \qquad (25)$$

where $\mathcal{A}_{\xi\eta}^{mn}$, $\mathcal{B}_{\xi\eta}^{mn}$, $\mathcal{C}_{\xi\eta}^{mn}$ and $\mathcal{D}_{\xi\eta}^{mn}$ are constants that are independent of the observation or source coordinates. Note that $\mathrm{Im}(k_{pzmn}) \geqslant 0$ and k_{pzmn} multiplies positive z-dependent quantities, bounding the absolute values of all exponentials in Eq. (25) between 0 and 1. Constants $\mathcal{A}_{\xi\eta}^{mn}$, $\mathcal{B}_{\xi\eta}^{mn}$, $\mathcal{C}_{\xi\eta}^{mn}$, and $\mathcal{D}_{\xi\eta}^{mn}$ can be derived from Eq. (6). For instance, one component is given below:

$$\mathcal{A}_{xx}^{mn} = -\frac{1}{2\epsilon_p D_x D_y} \cdot \frac{k_{pzmn}^2 k_{xm}^2 \widetilde{R}_{p,p-1}^{TM} + k_{yn}^2 k_p^2 \widetilde{R}_{p,p-1}^{TE}}{k_{pzmn}(k_p^2 - k_{pzmn}^2)} M_p, \qquad (26)$$

where $\widetilde{R}_{p,p-1}^{TM}$, $\widetilde{R}_{p,p-1}^{TE}$ and M_p depend on m and n. The integral Eq. (23) can be written as a sum

$$\mathcal{Z}_{ij}^{r} = i \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \mathcal{Z}_{Aij}^{mn} + \mathcal{Z}_{Bij}^{mn} + \mathcal{Z}_{Cij}^{mn} + \mathcal{Z}_{Dij}^{mn}, \qquad (27)$$

where

$$\mathcal{Z}_{Aij}^{mn} = \int_{V_i} d\mathbf{r}\, \mathbf{f}_i(\mathbf{r}) \cdot \int_{V_j} d\mathbf{r}'\, \mathcal{A}^{mn} \cdot \mathbf{f}_j(\mathbf{r}')$$

$$\times e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(z+z')]}, \qquad (28a)$$

$$\mathcal{Z}_{Bij}^{mn} = \int_{V_i} d\mathbf{r}\, \mathbf{f}_i(\mathbf{r}) \cdot \int_{V_j} d\mathbf{r}'\, \mathcal{B}^{mn} \cdot \mathbf{f}_j(\mathbf{r}')$$

$$\times e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(2h+z-z')]}, \qquad (28b)$$

$$\mathcal{Z}_{Cij}^{mn} = \int_{V_i} d\mathbf{r}\, \mathbf{f}_i(\mathbf{r}) \cdot \int_{V_j} 2d\mathbf{r}'\, \mathcal{C}^{mn} \cdot \mathbf{f}_j(\mathbf{r}')$$

$$\times e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(2h-z+z')]}, \qquad (28c)$$

$$\mathcal{Z}_{Dij}^{mn} = \int_{V_i} d\mathbf{r}\, \mathbf{f}_i(\mathbf{r}) \cdot \int_{V_j} d\mathbf{r}'\, \mathcal{D}^{mn} \cdot \mathbf{f}_j(\mathbf{r}')$$

$$\times e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(2h-z-z')]}. \qquad (28d)$$

Using the definition of the basis functions one can rewrite Eq. (28a) as

$$\mathcal{Z}_{Aij}^{mn} = s_i s_j (a_i a_j / 9 v_i v_j) \int_{V_i^{s_i}} d\mathbf{r}(\mathbf{r} - \mathbf{r}_i^{s_i}) \cdot \int_{V_j^{s_j}} d\mathbf{r}'\, \mathcal{A}^{mn}$$

$$\cdot (\mathbf{r}' - \mathbf{r}_j^{s_j}) e^{i[k_{xm}(x-x')+k_{yn}(y-y')+k_{pzmn}(z+z')]}, \qquad (29)$$

where the signs $\{s_i, s_j\} \in \{+1, -1\}$ are chosen according to the superscript of the corresponding tetrahedra $V = V_i^{s_i}$ and $V' = V_j^{s_j}$. Next, denote the uppermost vertex of tetrahedron V_i (the one with the highest z coordinate, or any of the uppermost vertices, if two or three vertices are situated equally high) as $\mathbf{r}_{up}$ and the uppermost vertex of tetrahedron V_j as $\mathbf{r}_{up}'$. Also, denote the lower vertices of V_i and V_j as $\mathbf{r}_{dn}$ and $\mathbf{r}_{dn}'$, respectively. Then, the double integral can be written as a vector-matrix-vector product with two independent integrals,

$$\mathcal{Z}_{Aij}^{mn} = s_i s_j \frac{a_i a_j}{9 v_i v_j} \int_{V_i^{s_i}} d\mathbf{r}[(\mathbf{r} - \mathbf{r}_{dn})$$

$$+ (\mathbf{r}_{dn} - \mathbf{r}_i^{s_i})] e^{i[k_{xm}(x-x_{dn})+k_{yn}(y-y_{dn})+k_{pzmn}(z-z_{dn})]}$$

$$\cdot \mathcal{A}^{mn} e^{i[k_{xm}(x_{dn}-x_{dn}')+k_{yn}(y_{dn}-y_{dn}')+k_{pzmn}(z_{dn}+z_{dn}')]}$$

$$\cdot \int_{V_j^{s_j}} d\mathbf{r}'[(\mathbf{r}' - \mathbf{r}_{dn}') + (\mathbf{r}_{dn}' - \mathbf{r}_j^{s_j})]$$

$$\times e^{i[k_{xm}(x_{dn}'-x')+k_{yn}(y_{dn}'-y')+k_{pzmn}(z'-z_{dn}')]}. \qquad (30)$$

In this integral, the choice of $\mathbf{r}_{dn}$ and $\mathbf{r}_{dn}'$ are critical. They are chosen such that the absolute value of all exponentials appearing in Eq. (30) are *bounded* between 0 and 1. This turns out to be crucial as it avoids roundoff errors due to multiplication of very large and very small numbers. Recalling that $\mathbf{k}_{mn} = k_{xm}\hat{x} + k_{yn}\hat{y} + k_{pzmn}\hat{z}$ and using the notations $\mathbf{k}_{mn}^{-} = k_{xm}\hat{x} + k_{yn}\hat{y} - k_{pzmn}\hat{z}$, $\mathbf{R}_{dn} = \mathbf{r} - \mathbf{r}_{dn}$ and $\mathbf{R}_{dn}' = \mathbf{r}' - \mathbf{r}_{dn}'$, the last expression can be written as

$$\mathcal{Z}_{Aij}^{mn} = \mathbf{I}_{dn} \cdot \mathbf{I}_{\mathcal{A}}, \tag{31}$$

where

$$\mathbf{I}_{dn} = \frac{s_i a_i}{3v_i} \left(\int_{V_i^{s_i}} d\mathbf{r} \ \mathbf{R}_{dn} e^{i\mathbf{k}_{mn} \cdot \mathbf{R}_{dn}} \right.$$
$$\left. + (\mathbf{r}_{dn} - \mathbf{r}_i^{s_i}) \int_{V_i^{s_i}} d\mathbf{r} \ e^{i\mathbf{k}_{mn} \cdot \mathbf{R}_{dn}} \right) e^{i\mathbf{k}_{mn} \cdot \mathbf{r}_{dn}}, \tag{32}$$

$$\mathbf{I}_{\mathcal{A}} = \mathcal{A}^{mn} \cdot \frac{s_j a_j}{3v_j} \left(\int_{V_j^{s_j}} d\mathbf{r}' \mathbf{R}_{dn}' e^{-i\mathbf{k}_{mn}^- \cdot \mathbf{R}_{dn}'} \right.$$
$$\left. + (\mathbf{r}_{dn}' - \mathbf{r}_j^{s_j}) \int_{V_j^{s_j}} d\mathbf{r}' e^{-i\mathbf{k}_{mn}^- \cdot \mathbf{R}_{dn}'} \right) e^{-i\mathbf{k}_{xm}^- \cdot \mathbf{r}_{dn}'}. \tag{33}$$

To cast the other three terms in Eq. (27) in a similar framework, they are written as

$$\mathcal{Z}_{Bij}^{mn} = \mathbf{I}_{dn} \cdot \mathbf{I}_{\mathcal{B}}, \tag{34}$$

$$\mathcal{Z}_{Cij}^{mn} = \mathbf{I}_{up} \cdot \mathbf{I}_{\mathcal{C}}, \tag{35}$$

$$\mathcal{Z}_{Dij}^{mn} = \mathbf{I}_{up} \cdot \mathbf{I}_{\mathcal{D}}, \tag{36}$$

where

$$\mathbf{I}_{up} = \frac{s_i a_i}{3v_i} \left(\int_{V_i^{s_i}} d\mathbf{r} \ \mathbf{R}_{up} e^{i\mathbf{k}_{mn}^- \cdot \mathbf{R}_{up}} \right.$$
$$\left. + (\mathbf{r}_{up} - \mathbf{r}_i^{s_i}) \int_{V_i^{s_i}} d\mathbf{r} \ e^{i\mathbf{k}_{mn} \cdot \mathbf{R}_{up}} \right) e^{i(k_{pzmn}h + \mathbf{k}_{mn}^- \cdot \mathbf{r}_{up})}, \tag{37}$$

$$\mathbf{I}_{\mathcal{B}} = \mathcal{B}^{mn} \cdot \frac{s_j a_j}{3v_j} \left(\int_{V_j^{s_j}} d\mathbf{r}' \mathbf{R}_{up}' e^{-i\mathbf{k}_{mn} \cdot \mathbf{R}_{up}'} \right.$$
$$\left. + (\mathbf{r}_{up}' - \mathbf{r}_j^{s_j}) \int_{V_j^{s_j}} d\mathbf{r}' e^{-i\mathbf{k}_{mn} \cdot \mathbf{R}_{up}'} \right) e^{i(2k_{pzmn}h - \mathbf{k}_{xm} \cdot \mathbf{r}_{up}')}, \tag{38}$$

$$\mathbf{I}_{\mathcal{C}} = \mathcal{C}^{mn} \cdot \frac{s_j a_j}{3v_j} \left(\int_{V_j^{s_j}} d\mathbf{r}' \mathbf{R}_{dn}' e^{-i\mathbf{k}_{mn}^- \cdot \mathbf{R}_{dn}'} \right.$$
$$\left. + (\mathbf{r}_{dn}' - \mathbf{r}_j^{s_j}) \int_{V_j^{s_j}} d\mathbf{r}' e^{-i\mathbf{k}_{mn}^- \cdot \mathbf{R}_{dn}'} \right) e^{i(k_{pzmn}h - \mathbf{k}_{xm}^- \cdot \mathbf{r}_{dn}')}, \tag{39}$$

$$\mathbf{I}_{\mathcal{D}} = \mathcal{D}^{mn} \cdot \frac{s_j a_j}{3v_j} \left(\int_{V_j^{s_j}} d\mathbf{r}' \mathbf{R}_{up}' e^{-i\mathbf{k}_{mn} \cdot \mathbf{R}_{up}'} \right.$$
$$\left. + (\mathbf{r}_{up}' - \mathbf{r}_j^{s_j}) \int_{V_j^{s_j}} d\mathbf{r}' e^{-i\mathbf{k}_{mn} \cdot \mathbf{R}_{up}'} \right) e^{i(k_{pzmn}h - \mathbf{k}_{xm} \cdot \mathbf{r}_{up}')}, \tag{40}$$

and $\mathbf{R}_{up} = \mathbf{r} - \mathbf{r}_{up}$, $\mathbf{R}_{up}' = \mathbf{r}' - \mathbf{r}_{up}'$.

It is apparent that in the above prescribed scheme, each of the integrals $\mathbf{I}_{dn}$, $\mathbf{I}_{up}$, $\mathbf{I}_{\mathcal{A}}$, $\mathbf{I}_{\mathcal{B}}$, $\mathbf{I}_{\mathcal{C}}$ and $\mathbf{I}_{\mathcal{D}}$ can be precomputed for every basis function with a cost that scales only as $\mathcal{O}(N)$ for every harmonic mn. These integrals are independent of test or source basis functions, and can be used to create an efficient framework for computing the entries of the matrix. The proposed technique also overcomes a significant hurdle in terms of convergence. If the inner product with the reflected Green's dyad were to be computed numerically, one would find that the reflected Green's dyad does not converge for source and observer pairs being co-located either on the top or the bottom surface of the film. As will be shown later, the proposed analytical technique overcomes this difficulty with aplomb. Care should be taken when evaluating these integrals as straightforward methods would lead to error growth when adding and subtracting large terms. The means to systematically avoid this is presented next.

2. Coordinate transformations for closed-form integrals

To compute each of the integrals in Eqs. (32), (33), and (37)–(40) it is convenient to transfer the origin of the coordinate system such that the vectors R_{dn}, R_{up}, R_{dn}', and R_{up}' are inside their respective domains of integration, i.e., inside tetrahedron $V_i^{s_i}$ for Eqs. (32) and (37) and inside tetrahedron $V_j^{s_j}$ for Eqs. (33) and (38)–(40), respectively. Specifically, if the origin of the coordinate system is transferred to $\mathbf{r}_{dn}$, then the test tetrahedron $V_i^{s_i}$ is moved to $V_i^{s_i} - \mathbf{r}_{dn}$ having one of its vertices at the origin. The two integrals in Eq. (32) become $\int_{(V_i^{s_i} - \mathbf{r}_{dn})} d\mathbf{r} \, \mathbf{r} e^{i\mathbf{k}_{mn} \cdot \mathbf{r}}$ and $\int_{(V_i^{s_i} - \mathbf{r}_{dn})} d\mathbf{r} \, e^{i\mathbf{k}_{mn} \cdot \mathbf{r}}$. On the other hand, if the origin of the coordinate system is transferred to $\mathbf{r}_{dn}'$, then the source tetrahedron $V_j^{s_j}$ is moved to $V_j^{s_j} - \mathbf{r}_{dn}'$ with one of its vertices at the origin. The two integrals in Eq. (33) become $\int_{(V_j^{s_j} - \mathbf{r}_{dn}')} d\mathbf{r} \, \mathbf{r} e^{-i\mathbf{k}_{mn}^- \cdot \mathbf{r}}$ and $\int_{(V_j^{s_j} - \mathbf{r}_{dn}')} d\mathbf{r} \, e^{-i\mathbf{k}_{mn}^- \cdot \mathbf{r}}$, respectively. Similarly, when computing Eqs. (37)–(40) the origin of the coordinate system can be moved to one of the points $\mathbf{r}_{up}$, $\mathbf{r}_{dn}'$, or $\mathbf{r}_{up}'$ so that only the two following general types of integrals I_1 and $\mathbf{I}_2$ are needed to get the closed-form analytical expression for Eq. (23),

$$I_1 = \int_V d\mathbf{r} \ e^{i\mathbf{k}^* \cdot \mathbf{r}}, \tag{41}$$

$$\mathbf{I}_2 = \int_V d\mathbf{r} \ \mathbf{r} e^{i\mathbf{k}^* \cdot \mathbf{r}}, \tag{42}$$

over tetrahedra having one vertex in the origin. In the above, the wave-number vector $\mathbf{k}^*$ can be one of the following four: $\mathbf{k}_{mn}$, $-\mathbf{k}_{mn}$, $\mathbf{k}_{mn}^-$, or $-\mathbf{k}_{mn}^-$, however, in any case the location of tetrahedra guarantees that all exponentials in the integrands are bounded in their absolute values between 0 and 1. This particular choice of vertices produces bounded exponentials necessary for avoiding cancellation errors.

To compute the integrals (41) and (42) with the origin at a vertex of a tetrahedron, it is convenient to use another coordinate system u, v, w as follows. Consider a tetrahedron $ABCO$ with O being the origin, as shown in Fig. 2. Let the w axis be normal to the plane ABC, with point $\mathbf{w}_0$ being the normal projection of point O onto the plane ABC. In what
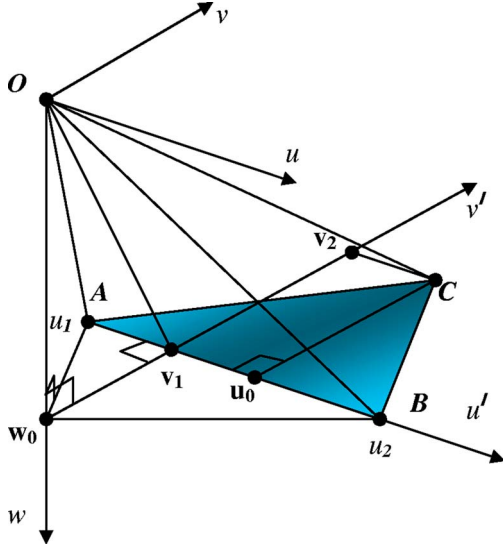
FIG. 2. (Color online) Local coordinate system for tetrahedron.

follows, we use $w_0 = |\overline{O\mathbf{w}_0}|$ and the unit vector $\hat{w} = \overline{O\mathbf{w}_0}/w_0$. The point $\mathbf{w}_0$ can be anywhere inside or outside of the triangle ABC. The quantities that are used to evaluate the integrals are defined as follows: $\mathbf{u}_0$ is the normal projection of C onto segment AB, $\hat{v} = \overline{\mathbf{u}_0C}/|\overline{\mathbf{u}_0C}|$ is the v' axis in the plane ABC normal to segment AB and pointing in the direction $\hat{v}$, $\mathbf{v}_1$ is the normal projection of $\mathbf{w}_0$ onto AB, $\mathbf{v}_2$ is the normal projection of C onto the v' axis, and the dot products $v_1 = \mathbf{w}_0\mathbf{v}_1 \cdot \hat{v}$ and $v_2 = \mathbf{w}_0\mathbf{v}_2 \cdot \hat{v}$. The u axis is chosen to be parallel to segment AB. Let

$$\left.\begin{array}{c} \mathbf{u}_1 = \overline{\mathbf{v}_1A} \\ \mathbf{u}_2 = \overline{\mathbf{v}_1B} \end{array}\right\} \quad \text{if } |\overline{\mathbf{v}_1B}| \geq |\overline{\mathbf{v}_1A}| \quad \text{or} \quad \left.\begin{array}{c} \mathbf{u}_1 = \overline{\mathbf{v}_1B} \\ \mathbf{u}_2 = \overline{\mathbf{v}_1A} \end{array}\right\} \quad \text{otherwise,}$$

$$(43)$$

$u_2 = |\mathbf{u}_2|$, $\hat{u} = \mathbf{u}_2/u_2$, and $u_1 = \mathbf{u}_1 \cdot \hat{u}$. Then an arbitrary point in the new coordinate system is $\mathbf{r} = u\hat{u} + v\hat{v} + w\hat{w}$, where $u = \mathbf{r} \cdot \hat{u}$, $v = \mathbf{r} \cdot \hat{v}$, and $w = \mathbf{r} \cdot \hat{w}$. Taking the dot products of the wave-number vector $k_u = \mathbf{k}^* \cdot \hat{u}$, $k_v = \mathbf{k}^* \cdot \hat{v}$, and $k_w = \mathbf{k}^* \cdot \hat{w}$, integral (41) is calculated as

$$I_1 = \int_0^{w_0} dw \int_{wv_{w1}}^{wv_{w2}} dv \int_{(wu_{w1}+vu_{v1})}^{(wu_{w2}+vu_{v2})} e^{i(k_u u + k_v v + k_w w)} du, \quad (44)$$

where $v_{w1} = v_1/w_0$, $v_{w2} = v_2/w_0$, $u_{w1} = u_1/w_0 - v_1(u_0 - u_1)/w_0(v_2 - v_1)$, $u_{w2} = u_2/w_0 - v_1(u_0 - u_2)/w_0(v_2 - v_1)$, $u_{v1} = (u_0 - u_1)/(v_2 - v_1)$ and $u_{v2} = (u_0 - u_2)/(v_2 - v_1)$. Similarly the vector integral (42) is written as

$$\mathbf{I}_2 = \hat{u}I_u + \hat{v}I_v + \hat{w}I_w, \quad (45a)$$

$$I_u = \hat{u} \int_0^{w_0} dw \int_{wv_{w1}}^{wv_{w2}} dv \int_{(wu_{w1}+vu_{v1})}^{(wu_{w2}+vu_{v2})} e^{i(k_u u + k_v v + k_w w)} u\, du,$$

$$(45b)$$

$$I_v = \hat{v} \int_0^{w_0} dw \int_{wv_{w1}}^{wv_{w2}} v\, dv \int_{(wu_{w1}+vu_{v1})}^{(wu_{w2}+vu_{v2})} e^{i(k_u u + k_v v + k_w w)} du,$$

$$(45c)$$

$$I_w = \hat{w} \int_0^{w_0} w\, dw \int_{wv_{w1}}^{wv_{w2}} dv \int_{(wu_{w1}+vu_{v1})}^{(wu_{w2}+vu_{v2})} e^{i(k_u u + k_v v + k_w w)} du.$$

$$(45d)$$

The first of the above four integrals in Eqs. (44) and (45b)–(45d) can be evaluated in closed form as

$$I_1 = \frac{iw_0\Delta v}{k_u} \left[\frac{1}{\psi_1}\left(\frac{e^{i\phi_1}-1}{\phi_1} - \frac{e^{i\phi_0}-1}{\phi_0} \right) - \frac{1}{\psi_2}\left(\frac{e^{i\phi_2}-1}{\phi_2} - \frac{e^{i\phi_0}-1}{\phi_0} \right) \right], \quad (46)$$

where we defined $\Delta v = v_2 - v_1$, $\psi_1 = \phi_v + k_u(u_0 - u_1)$, $\psi_2 = \phi_v + k_u(u_0 - u_2)$, $\phi_1 = \phi_w + k_v v_1 + k_u u_1$, $\phi_0 = \phi_w + k_v v_2 + k_u u_0$, and $\phi_2 = \phi_w + k_v v_1 + k_u u_2$, $\phi_v = k_v\Delta v$, $\phi_w = k_w w_0$. Expressions for the integrals I_u, I_v, and I_w are provided in Appendix B. This expression is not singular provided all the elements in the set $\Psi = \{k_u, \psi_1, \psi_2, \phi_0, \phi_1, \phi_2\}$ are not zeros, i.e.,

$$\phi \neq 0 \; \forall \; \phi \in \Psi. \quad (47)$$

If this condition is not met, another closed form should be chosen, as elucidated in detail in Appendix B. Similarly, it is shown there that the integrals to compute Eqs. (45b)–(45d) have several expressions each of which must be used for a particular combination of nonzero subsets in Ψ.

V. RESULTS

In what follows, we shall present three sets of results. The first will serve to validate the accuracy of the analytical scheme presented in the last section. With these we will demonstrate that the use of analytical integration is imperative if one is to accurately simulate fields scattered by periodic apertures in multilayered structures. Next, we will validate the results generated by the methodology presented herein against both analytical data and complementary structures. Finally, we will demonstrate that the proposed technique accurately predicts the location of enhanced transmission for a couple of configurations.

A. Analytical vs numerical integration

To validate the correctness and usefulness of the analytical expressions for the reflected part of the Green's function, the integrals (23) have been computed analytically and using Gaussian quadrature. This comparison has been done for all possible combinations of i and j in a small but representative case. Specifically, holes have been discretized with 40 tetrahedra with $h = 12.5$ nm, $W_x = W_y = 9.375$ nm, and $D_x = D_y = 581.7$ nm. A plane wave is assumed to be normally incident on the structure and has a wavelength of $\lambda_0 = 450$ nm. The integral over a spectral mode is evaluated both analytically and numerically using four quadrature rules; first-, second-,
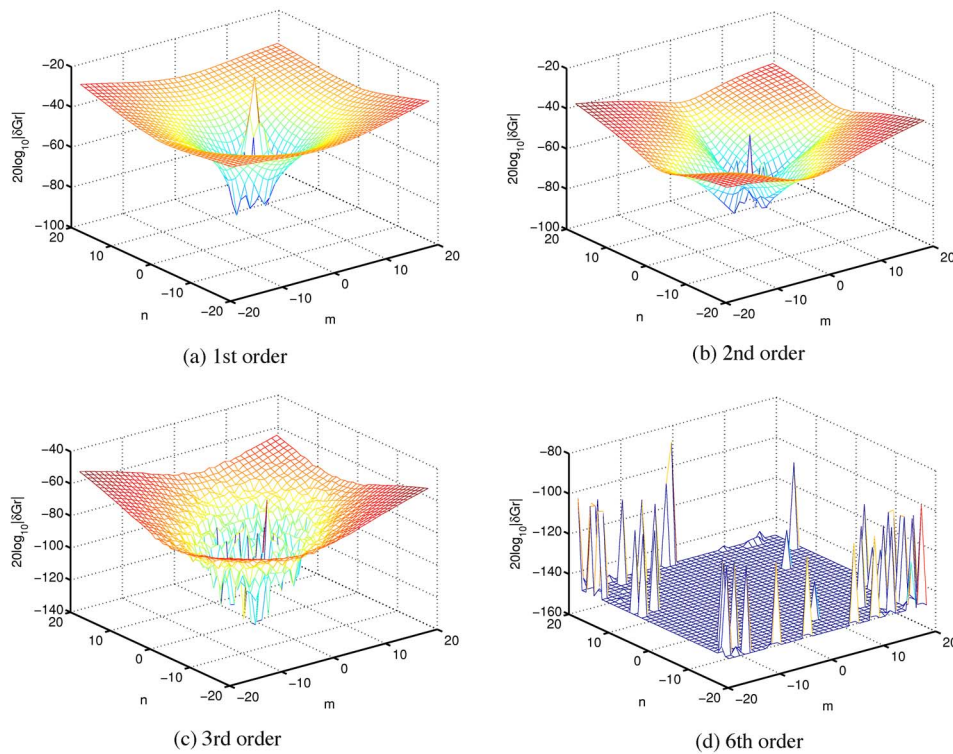
(a) 1st order

(b) 2nd order

FIG. 3. (Color online) Relative absolute error of spectral components of integral (23) for different order Gaussian quadrature rules (a) first order; (b) second order; (c) third order; (d) sixth order.

(c) 3rd order

(d) 6th order

third-, and sixth-order Gaussian quadratures. These correspond to 1-pt, 4-pt, 5-pt, and 25-pt integration rules. In all cases, the relative error $20\log_{10}|\mathcal{Z}^{mn}_{ij\text{ANALYTICAL}} - \mathcal{Z}^{mn}_{ij\text{NUMERICAL}}|/|\mathcal{Z}^{mn}_{ij\text{ANALYTICAL}}|$ is plotted vs mode numbers $n,m=\pm20$. The pair chosen for demonstration is random—it was tested that the conclusions of this study hold for all pairs. Also, it requires $\{\max(N_F),M_F\}=1200$ modes for the analytical integral to converge to a relative error of 10^{-4}. These data were obtained using the analytical integration rules. As is expected, the first order integration rule is not particularly good as is seen in Fig. 3(a). The minimum error is -50 db, and the error over the higher order modes increases rapidly. Figure 3(b) shows the relative error when the integrals (23) are computed numerically using second order Gaussian quadrature rule. It is seen that the components around central spatial frequency are computed very accurately (error around -90 dB), however, higher spectral components result in considerably higher errors (around -30 dB at $n,m=\pm20$). Exhaustive analysis of all possible pairs $\{i,j\}$ showed similar error plots with relative error ranging from -33 dB to -27 dB at $n,m=\pm20$. Thus, given the nature of Fig. 3(b), the error becomes unacceptable as higher order spectral modes are involved. Better results can be achieved when third-order Gaussian integration rule is used. Figure 3(c) shows that in this case the relative error is smaller and is around -50 dB to -47 dB at $n,m=\pm20$. However, it is evident from the nature of the figure that errors increase with increase in the mode number, and the accumulated error will be high. Finally, Fig. 3(d) shows the relative error for the sixth order Gaussian quadrature rule with 25 integration points per tetrahedron (or 625 pairs per element in $\mathcal{Z}$). For mode numbers $-20 \leqslant m,n \leqslant +20m$ the spectral components are computed with error below -90 dB (most of the compo-

nents below -200 dB). These data indicate two salient facts: (i) the numerical quadratures converge to analytical results over the *entire* range of the necessary mode numbers *only* for very high quadrature rules; (ii) using these rules would make the evaluation of the inner products using numerical quadratures very expensive. Thus it is imperative that one use analytical techniques prescribed herein when evaluating the inner products. For example, the time spent to compute $\mathcal{Z}^r_{ij}$ analytically in Eq. (23) on an Ultra-SPARC III 750 MHz 64-bit processor was 103 s, and 184 s, 563 s, 776 s and 10 516 s for the first, second, third, and sixth order quadratures, respectively. It should also be noted that if one were to use numerical quadrature to compute all elements of $\mathcal{Z}^r$, one would need to evaluate the nested double integrals for all pairwise integrations. In contrast, the method presented here permits the precomputing of the integrals over the domain of the basis function for each mode, and matrix elements are obtained by simply translation and addition.

Costs associated with the proposed scheme

It has been shown thus far that numerical integration converges to the analytical result albeit at considerable cost in terms of CPU time. Aside from this fact, the time for computation of the matrix scales as $\mathcal{O}(N)$ per harmonic (nm), where N is the number of basis functions chosen to represent the flux density. The number of basis functions is chosen such that the edge length of the basis function is approximately one-fourth of the skin depth. However, the key to rapid computation of the number of harmonics that it takes for the matrix to converge is the presence of the term $k_z f(z,z')$. Here, $f(z,z')$ is a function of the z coordinates of the source and observer locations. For materials being analyzed here, this always has a negative real part, and for non-
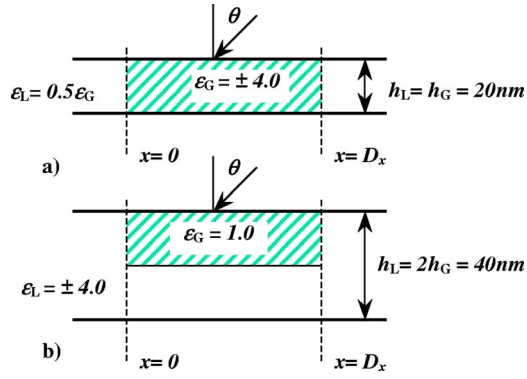
FIG. 4. (Color online) Slab models: (a) inclusion material G occupies all second layer, and (b) partially substitutes layer.

FIG. 5. (Color online) Transmitted power vs incidence angle θ at free-space wavelength $\lambda_0 = 400$ nm for 20 nm slab: TE (solid lines) and TM (dashed lines) modes computed analytically are compared with nu- merical solution when second layer completely substituted with inclusions of $\epsilon_G = 2\epsilon_L$ (cross marks) and when second layer is partially removed (circles). Dots represent classical volume equivalence principle formulation results.

zero $f(z,z')$ this factor will contribute to exponential convergence of the spectral Green's function. Indeed, our tests have shown that it may take as few as six spectral harmonics for the Green's function to converge to a relative error of 10^{-4}. The problem, however, is when $f(z,z')$ approaches zero. This typically occurs when both the source and observer pairs are at either the top or bottom surfaces. While it considerably mitigates the problems caused by $f(z,z')$ approaching zero, the convergence is nonetheless affected. As mentioned earlier, the maximum number of harmonics required is of the order of 1200 to converge to a relative error of 10^{-4}. Thus, while the cost of filling the matrix has been reduced to $\mathcal{O}(N)$ from $\mathcal{O}(N^2)$, the solve time still scales as $\mathcal{O}(N^2)$. As the matrix is full, the memory requirements scale as $\mathcal{O}(N^2)$.

B. Validation of the proposed secheme

To validate the algorithm, an exhaustive set of special cases has been tested. First, a variety of stratified layered media have been checked. Reflected and transmitted power have been computed and compared against exact analytical data. In the first set of examples, a three-layered medium is considered. The first and third layers are free space and the second layer is $h = 20$ nm thick and of relative permittivity ε_L. The perforation is of dimension $W_x \times W_y$ and is made of a material of relative permittivity ε_G. It is assumed that the wavelength of the incident wave is 400 nm, and the angles of incidence range from 0° to 90°. To validate our code against analytical data, we let the periodicity be equal to the width of the protrusion, i.e., $D_x = W_x$ and $D_y = W_y$. Physically, this is tantamount to replacing the second layer with one whose permittivity is ε_G. The reflection and transmission coefficients for this geometric configuration can be computed analytically. Note that in the numerical scheme prescribed in the paper, this is done via equivalent currents. Different combinations of ε_L and ε_G were chosen, and results were obtained for both TE and TM polarizations. More specifically, the perforation geometry and periodicity are related via $D_x = D_y = W_x = W_y = 20$ nm and the geometry was discretized using 480 tetrahedra. Both positive and negative permittivity have been tested. Figure 4(a) shows the configuration of the problem and Fig. 5 shows the transmitted power vs incident angle θ for TM and TE cases. The solid lines correspond to TE
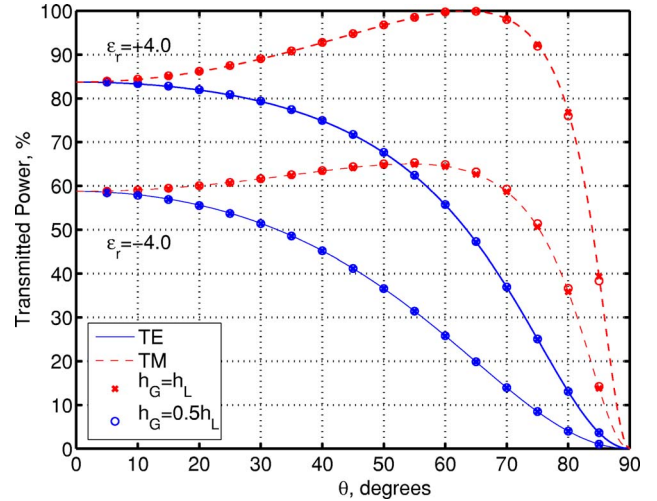
exact solutions when $\epsilon_G = 4.0$ and when $\epsilon_G = -4.0$, and the dashed lines correspond to TM exact solutions. In these figures, $\times$ correspond to the cases when $\epsilon_L = 0.5\epsilon_G$ and the dots correspond to the cases with $\epsilon_L = 1.0$. As is evident from these figures, the agreement with analytical data is excellent.

Next, we analyze a three layer example wherein the thickness of the later is $h = 40$ nm. The height of the perforation is 20 nm, and it is flush with the top surface of the film. If the periodicity is made identical to the width and breadth of the protrusion, then this is tantamount to replacing half the film with a material whose relative permittivity is ε_G. The analytical results for such a configuration can be computed analytically. Figure 4(b) shows the geometric configuration for this problem. For this numerical experiment, the dimensions of the protrusion, the periodicity, and the discretization were chosen to be the same as before. Also, $\varepsilon_G = 1.0$ and $\varepsilon_L = \pm 4.0$. The transmitted power as a function of the incident angle θ for TM and TE cases is shown in Fig. 5. It is seen that for a variety of configurations, corresponding to a slab of constant permittivity the transmission has been calculated very accurately.

Thus far, we have presented comparison against analytical data. Next, we analyze a structure using two complementary methods. The objective is to compute reflection and transmission from rods of rectangular cross section that are positioned parallel to the x axis and periodic along the y axis. The width and height of the rods are $W_x = h = 20$ nm. The relative permittivity of the structure is ± 4.0. Figure 6(a) depicts the structure being modeled. Two different equivalence theorems can be used to solve the problem: (i) volume equivalent currents are used to replace the rods, and these currents radiate in the presence of free space; (ii) volume equivalent currents are used to represent the free space between the rods, and these radiate in the presence of layered media. These equivalences are depicted in Figs. 6(b) and
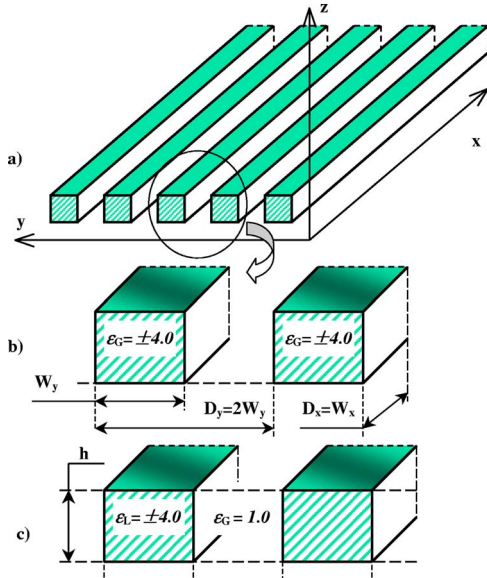
FIG. 6. (Color online) Complementary model types for a structure of stacked rods (a) when periodic volume-in-free-equivalence principle is used (b) and for the periodic layered medium formulation (c).

6(c), respectively. While both formulations invoke periodicity, it is the characteristics of the Green's functions that make the methods of analysis significantly different—in the former case, the reflected Green's dyad is zero, thereby considerably simplifying the analysis. To analyze this structure for both TE and TM polarizations and for a range of incidence angles, the rectangular prism of dimension $20 \times 20 \times 20$ nm^3 is discretized using 480 tetrahedra. The period for the analysis is $D_x = 2W_x = 40$ nm and $D_y = W_y = 20$ nm. In the first case, $\varepsilon_L = 1$ and $\varepsilon_G = \pm 4.0$, and in the second case $\varepsilon_L = \pm 4.0$ and $\varepsilon_G = 1.0$. Figure 7 shows the transmitted power for positive and negative permittivity, for x-polarized and y-polarized inci-
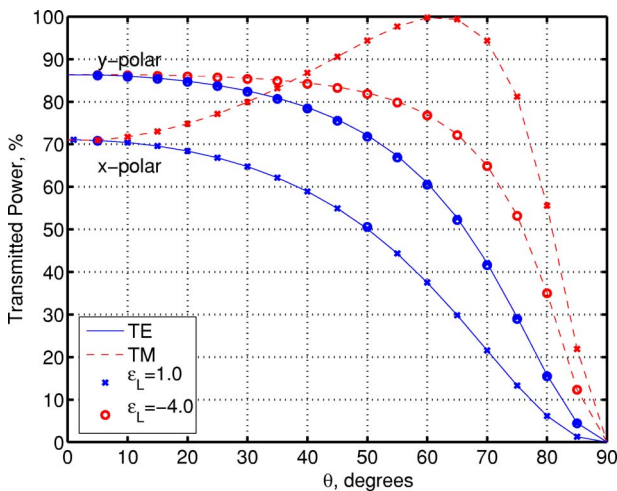
FIG. 7. (Color online) Transmitted power vs incidence angle θ at free-space wavelength $\lambda_0 = 400$ nm for rod stack structure computed as dielectric in free space ("×" marks) and as free-space inclusions in dielectric layer (circles). Solid lines connect TE and dashed lines connect TM solutions.
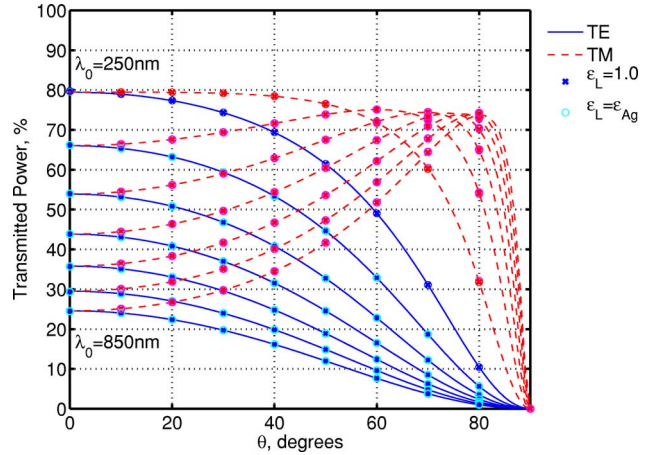

FIG. 8. (Color online) Transmitted power vs incidence angle θ at various frequecies of TE and TM excitation for silver 12.5 nm thick film. Analytically computed TE (solid lines) and TM (dashed lines) solution is compared with formulation in free space (cross marks) and periodic layered medium formulation (circles).

dent electric field, for both TE and TM modes. The transmitted power computed for the first method is represented using crosses, and the transmitted power for the second method is shown with circles. The solid lines connecting the marks show the TE mode, whereas the dashed lines represent the TM mode. As is evident from this figure, the results for the two complementary modeling methods are virtually identical, i.e., the performance of periodic layered Green's functions in the proposed numerical scheme has been validated with homogeneous periodic Green's functions used in the classical volume equivalence method.

The above experiments show that the proposed method performs exceptionally well for lossless dielectric layers whose permittivity is either positive or negative. As our goal is to develop a technique for the analysis of noble materials, it is imperative that one uses models that accurately capture
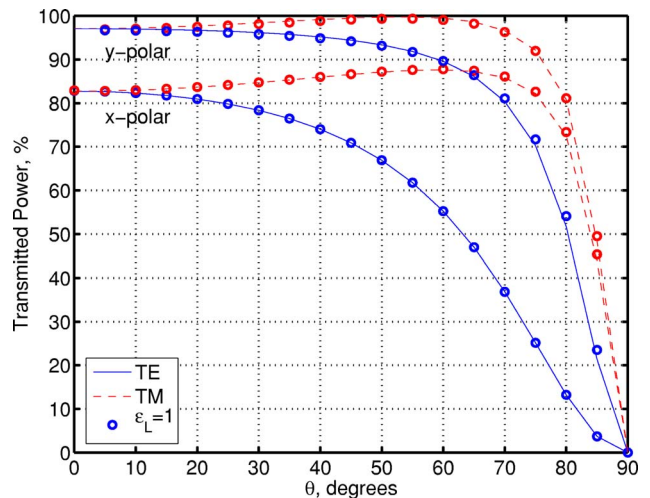

FIG. 9. (Color online) Transmitted power vs incidence angle θ at $\lambda_0 = 450$ nm for complementary structure with silver. Cross marks show free-space formulation, circles show periodic layered medium formulation results.
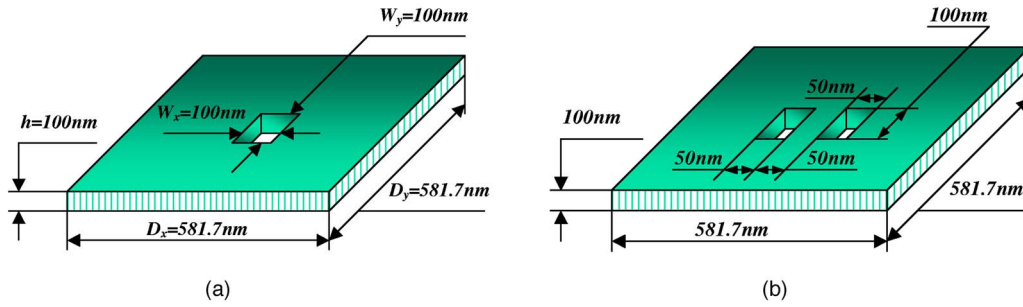
FIG. 10. (Color online) Two geometric configurations being studied: (a) a rectangular perforation; (b) two slots next to each other.

the frequency response of the dielectric. In the formulation presented herein it is trivial to incorporate this behavior and we do so by using the Drude free-electron model. The following parameters are used to describe the relative permittivity of silver:

$$\varepsilon_{Ag} = 1 + \frac{i\tau\omega_p^2}{2\omega(1 - i\omega\tau)} \tag{48}$$

with $\omega = 2\pi f$ the wave frequency, $\omega_p = 1.32 \times 10^{16}$ s^{-1} the plasma frequency, $\tau = 1.45 \times 10^{-14}$ s the plasma relaxation time. Figure 8 compares the numerically computed and the analytical transmitted power vs the angle of incidence θ for a 12.5-nm-thick silver slab. The discretization of the structure is exactly the same as was used earlier. The data are obtained for a range of free-space wavelengths starting from $\lambda_0 = 250$ nm and longer. The calculated transmission, denoted by circles, is compared with exact analytical results for TE (solid lines) and TM (dashed lines) excitation modes. We have also computed the same data using the classical volume equivalence theorem (currents radiate in free space), and these results are depicted as crosses in the figure. As is evident from the figure, the agreement between analytical and numerical data is very good. This is true over a wide range of frequencies wherein the real part of the permittivity varies from −2.07 at $\lambda_0 = 250$ nm to −34.45 at $\lambda_0 = 850$ nm. The variation in skin depth over this range of frequencies is not that spectacular—it varies from 27.7 to 23.0 nm.

Finally, we analyze scattering from periodically arranged silver rods using two complementary approaches as was done earlier. In this case, the specifications are as follows:

the rectangular cross section of the rods is $W_x = W_y = 9.375$ nm, $h = 12.5$ nm high, and the periods in the x and y directions are chosen to be $D_x = W_x$ and $D_y = 2W_y$, respectively. The geometry was discretized with 480 tetrahedra, and the transmitted power for such a structure was computed for incident angles ranging from $\theta = 0°$ to $90°$. Both TE and TM polarizations were tested at $\lambda_0 = 450$ nm. Figure 9 compares the results obtained using the proposed technique with those obtained using the classical volume equivalence theorem, and they agree well with each other.

C. Structures with periodical inclusions

The next three sets of numerical experiments analyze scattering from perforated films. In all cases, the thickness of the film is 100 nm, and the periods in the $\hat{x}$ and $\hat{y}$ directions are $D_x = D_y = 581.7$ nm, respectively. In all cases, transmission and reflection are computed over a range of frequencies for two cases: (i) using $\varepsilon_p = \mathrm{Re}\{\varepsilon_{Ag}\}$, and (ii) $\varepsilon_p = \varepsilon Ag$, i.e., we compute the transmission with and without loss. The field is assumed to be normally incident to the surface. For all cases, using the distribution of the real part of the relative permittivity of the film with frequency, and empirical formulas derived for the wave number at which the surface plasmon is excited, we expect enhanced transmission around 600 nm. The first example is a square perforation of side length ($W_x = W_y = 100$ nm) as shown in Fig. 10(a). The transmission coefficients are shown for both cases in Fig. 11. Indeed, when only the real part is used, near total transmission can be observed at 609 nm, and a narrower but sharp enhancement at 602.5 nm. However, when losses are included in the struc-
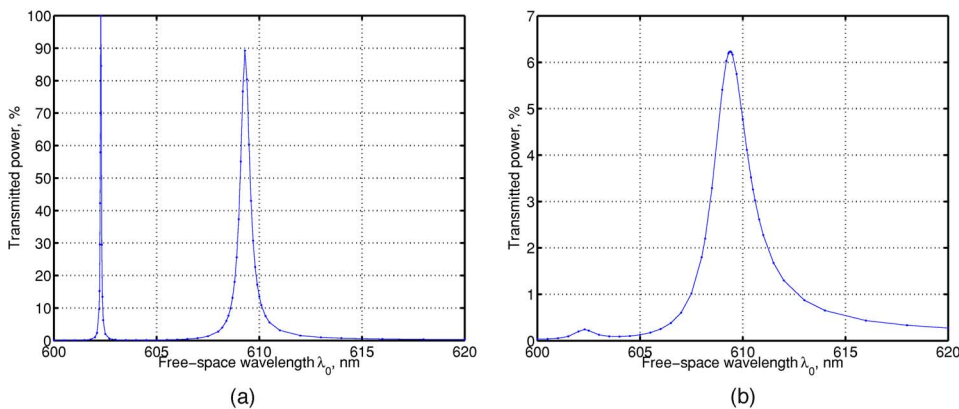

FIG. 11. (Color online) Transmitted power for normally incident wave impinging on 100 nm thick material perforated with 100 nm square holes periodically. Period is 581.7 nm in x and y directions. (a) $\varepsilon_p = \mathrm{Re}\{\varepsilon_{Ag}\}$, (b) $\varepsilon_p = \varepsilon_{Ag}$.
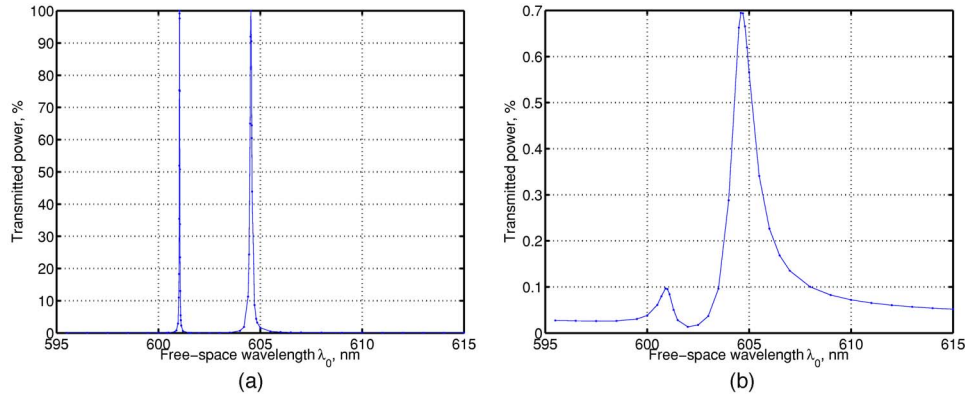
FIG. 12. (Color online) Transmitted power for normally incident wave impinging on 100 nm thick material perforated periodically with two 50×100 nm slots 50 nm apart. Period is 581.7 nm in x and y directions, and the incident field is $\hat{x}$ polarized. (a) $\varepsilon_p = \mathrm{Re}\{\varepsilon_{Ag}\}$, (b) $\varepsilon_p = \varepsilon_{Ag}$.

ture, both peaks are considerably smaller, and the transmitted power amounts to only 6%. Parenthetically, we note that the perforations studied here are smaller than those studied elsewhere.

Next, we study a more complex structure as shown in Fig. 10(b). In the first test the incident field is polarized along the $\hat{x}$ axis. Again as shown in Fig. 12(a), enhanced transmission is observed at 601.5 and 604.25 nm for the lossless case. The peaks are narrower than before. The inclusion of loss has a rather dramatic effect on the transmission coefficient as seen in Fig. 12(b)—while the peaks are visible, the magnitude is considerably smaller. Finally, we assume that the incident field is polarized in the $\hat{y}$ direction. Again, for the lossless case, enhanced transmission peaks are observed at 602.5 and 610 nm as seen in Fig. 13(a)—remarkably close to that observed for Fig. 10(a). This is readily explained by the coincidence of the polarization and the principal direction of the slots. However, inclusion of loss considerably decreases the transmitted field as seen in Fig. 13(b).

What is immediately apparent from the above figures is the nature of the enhanced transmission that is observed—it is a consequence of strong coupling to the apertures on either side of the film which provides the necessary mechanism for the tunneling of the fields [56,57]. Indeed, results similar to those in Figs. 11(a), 12(a), and 13(a) can be obtained for perfectly conducting structures as well. This, of course, has very little to do with SPPs, and is intimately linked to the coupling of surface waves on the periodic structure with the waveguide modes that exist in the slots. However, at optical

frequencies, one necessarily needs to include the proper material behavior. This entails the inclusion of loss in the simulation. Field penetration within the structure considerably affects the total transmission. Indeed, as is evident in Figs. 11(b), 12(b), and 13(b), the transmitted peak is considerably smaller, and more dispersed. Thus designs that rely only on using only the real part of the permittivity will be incorrect. Indeed, the results obtained show similar behavior (in terms of order of magnitude) to those obtained experimentally by Ghaemi *et al.* [58].

VI. SUMMARY

In this paper, we have developed the means to study scattering from perforated metallic films. The prescribed technique is both accurate and efficient. As is evident from the results, it can be used for the entire spectrum of incident angles—even at almost grazing. The technique has been extensively validated both analytical data as well as against other numerical methods. While the formulation that we have used relies on a volume integral equation, it can very well be used in a surface integral setting as well. The only problem with the latter is the slow convergence of the reflected Green's dyad. It is evident that portions of the techniques presented in this paper may be readily ported to the analysis of other applications—specifically, the means to accelerate the Ewald summation and analytical integration of the spectral Green's function. Our current work focuses on the application of this technique to a range of configurations of perforations, and extending the methods presented
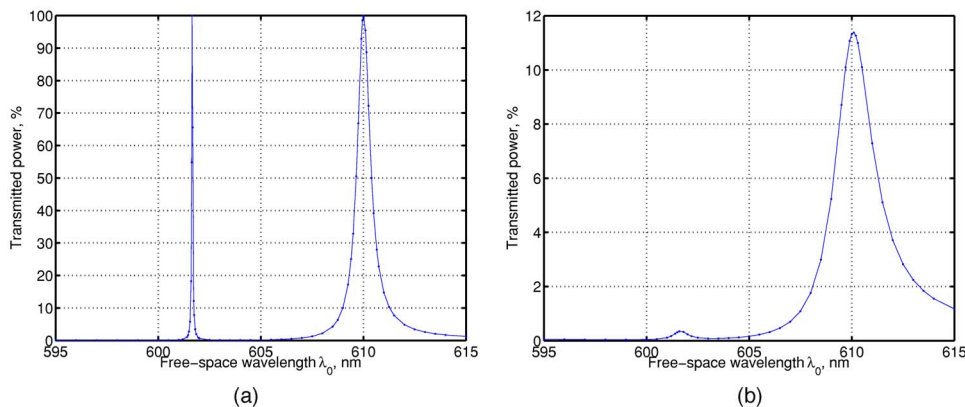

FIG. 13. (Color online) Transmitted power for normally incident wave impinging on 100 nm thick material perforated periodically with two 50×100 nm slots 50 nm apart. Period is 581.7 nm in x and y directions, and the incident field is $\hat{y}$ polarized. (a) $\varepsilon_p = \mathrm{Re}\{\varepsilon_{Ag}\}$, (b) $\varepsilon_p = \varepsilon_{Ag}$.

herein to time domain. These results will be presented elsewhere.

APPENDIX A: SEPARATION OF THE LAYERED MEDIUM GREEN'S FUNCTION INTO DIRECT AND REFLECTED PARTS

This appendix provides the formulas for expressing the Green's dyad for a layered medium as a sum of the directed and reflected portions. As noted earlier, such a partitioning is essential as the direct portion contains a source point singularity and has to be dealt with in a special manner. Once this separation is achieved, it is relatively straightforward to transform these results into their periodic counterparts. While some details can be found in Ref. [42], a more complete derivation can be found in Ref. [59]. The configuration being analyzed is as follows: Consider a M layer body that contains an electric dipole $\mathbf{J}(\mathbf{r}') = Il\hat{a}\delta(\mathbf{r}')$ in the pth layer whose thickness is denoted by h. It is assumed that the observation point also lies in the same layer. The Green's dyad can be written as

$$\mathcal{G}(\mathbf{r},\mathbf{r}') = i\omega\left(\mathcal{I} + \frac{1}{k_p^2}\nabla\nabla\right)g(|\mathbf{r}-\mathbf{r}'|) + \mathcal{G}_r(\mathbf{r},\mathbf{r}')$$

$$= \mathcal{G}_d(\mathbf{r},\mathbf{r}') + \mathcal{G}_r(\mathbf{r},\mathbf{r}'). \tag{A1}$$

The reflected Green's dyad is given by

$$\mathcal{G}_{xx}^r = -\int\int_{-\infty}^{\infty}dk_xdk_ye_{xy}\frac{k_{zp}^2k_x^2F_r^{TM}(z,z')+k_y^2k_p^2F_r^{TE}(z,z')}{k_\rho^2},$$

$$\mathcal{G}_{xy}^r = -\int\int_{-\infty}^{\infty}dk_xdk_ye_{xy}k_xk_y\frac{k_{zp}^2F_r^{TM}(z,z')-k_p^2F_r^{TE}(z,z')}{k_\rho^2},$$

$$\mathcal{G}_{xz}^r = -\int\int_{-\infty}^{\infty}dk_xdk_ye_{xy}k_xk_{zp}dF_r^{TM}(z,z'),$$

$$\mathcal{G}_{yx}^r = \mathcal{G}_{xy}^r,$$

$$\mathcal{G}_{yy}^r = -\int\int_{-\infty}^{\infty}dk_xdk_ye_{xy}\frac{k_{zp}^2k_y^2F_r^{TM}(z,z')+k_x^2k_p^2F_r^{TE}(z,z')}{k_\rho^2},$$

$$\mathcal{G}_{yz}^r = -\int\int_{-\infty}^{\infty}dk_xdk_ye_{xy}k_yk_{zp}dF_r^{TM}(z,z'),$$

$$\mathcal{G}_{zx}^r = -\mathcal{G}_{xz}^r,$$

$$\mathcal{G}_{zy}^r = -\mathcal{G}_{yz}^r,$$

$$\mathcal{G}_{zz}^r = -\int\int_{-\infty}^{\infty}dk_xdk_ye_{xy}k_\rho^2F_r^{TM}(z,z'), \tag{A2a}$$

where

$$F_r^\gamma(z,z') = \left[\widetilde{R}_{p,p+1}^\gamma e^{ik_{zp}(2h-z-z')} + \widetilde{R}_{p,p-1}^\gamma e^{ik_{zp}(z+z')}\right.$$
$$\left. + \widetilde{R}_{p,p-1}^\gamma\widetilde{R}_{p,p+1}^\gamma\left(e^{ik_{zp}(2h-z+z')} + e^{ik_{zp}(2h+z-z')}\right)\right]M_p^\gamma, \tag{A3a}$$

$$dF_r^\gamma(z,z') = \left[\widetilde{R}_{p,p+1}^\gamma e^{ik_{zp}(2h-z-z')} - \widetilde{R}_{p,p-1}^\gamma e^{ik_{zp}(z+z')}\right.$$
$$\left. + \widetilde{R}_{p,p-1}^\gamma\widetilde{R}_{p,p+1}^\gamma\left(e^{ik_{zp}(2h-z+z')} - e^{ik_{zp}(2h+z-z')}\right)\right]M_p, \tag{A3b}$$

$$M_p^\gamma = \left[1 - \widetilde{R}_{p,p-1}^\gamma\widetilde{R}_{p,p+1}^\gamma e^{i2k_{zp}h}\right]^{-1}, \tag{A3c}$$

and $\widetilde{R}_{p,p-1}^\gamma$, $\widetilde{R}_{p,p+1}^\gamma$ are the generalized reflection coefficients for $\gamma=$TE/TM polarization. The generalized TM and TE reflection coefficient can be easily found recursively [42,43]. For three layers and both the source and observer being in the second layer,

$$R_{p,p+1}^{TE} = \frac{\mu_{p+1}k_{zp} - \mu_pk_{zp+1}}{\mu_{p+1}k_{zp} + \mu_pk_{zp+1}},$$

$$R_{p,p+1}^{TM} = \frac{\epsilon_{p+1}k_{zp} - \epsilon_pk_{zp+1}}{\epsilon_{p+1}k_{zp} + \epsilon_pk_{zp+1}}. \tag{A4}$$

Note that R^{TE} is the reflection coefficient for the electric field and R^{TM} is the reflection coefficient for the magnetic field.

Remark: Note that the relations (A2d), (A2g), and (A2h) reflect the reciprocity theorem for the layered medium. For example, the spatial domain reciprocity $E_{xz}(\mathbf{r},\mathbf{r}') = E_{zx}(\mathbf{r}',\mathbf{r})$ translates into its spectral domain counterpart as $E_{xz}(k_x,k_y,z,z') = E_{zx}(-k_x,-k_y,z',z)$. This manifests itself in relations $k_xe_0\,\text{sgn}(z-z') = -k_xe_0\,\text{sgn}(z'-z)$ for the principal part and $k_xe_0dF_r^{TM}(z,z') = -k_xe_0dF_r^{TM}(z',z)$.

1. Transmission and reflection

To find the total transmission one needs to find the field in the last layer M. Only the upgoing wave exists, i.e., for $z > d_{M-1}$,

$$F_r^\gamma(z,z') = B_M^\gamma e^{ik_{zM}z}, \tag{A5a}$$

$$\frac{\partial F_r^\gamma(z,z')}{\partial z} = ik_{zM}F_r^\gamma(z,z'), \tag{A5b}$$

and B_M^γ is found as

$$B_M^\gamma = \widetilde{T}_{p,M}^\gamma e^{-i(k_{zM}d_{M-1} - \sum_{s=p+1}^{M-1}k_{zs}h_s)}\left(B_p^\gamma e^{ik_{zp}h} + e^{ik_{zp}(h-z')}\right), \tag{A6}$$

where $\widetilde{T}_{s,s+1}^\gamma$ is the generalized transmission coefficient from layer s to layer $s+1$, and the notation $\widetilde{T}_{p,M}^\gamma = \prod_{s=p}^{M-1}\widetilde{T}_{s,s+1}^\gamma$ has

been used. To find the total reflection the field in layer 1 is needed. Only the downward propagation exists, i.e., for $z < d_1$,

$$F_r^\gamma(z,z') = D_1^\gamma e^{-ik_{z1}z}, \tag{A7a}$$

$$\frac{\partial}{\partial z} F_r^\gamma(z,z') = -ik_{z1} F_r^\gamma(z,z'), \tag{A7b}$$

and D_1^γ is found as

$$D_1^\gamma = \tilde{T}_{p,1}^\gamma (D_p^\gamma + e^{ik_{zp}z'}) e^{i(k_{z1}d_1 + \Sigma_{s=2}^{p-1} k_{zs}h_s)}, \tag{A8}$$

where $\tilde{T}_{p,1}^\gamma = \Pi_{s=1}^{p-1}\tilde{T}_{s+1,s}^\gamma$. Transmitted and reflected electric field now can be computed using Eq. (A2). The magnetic field spectral components can be found from the electric field spectral components as $\mathbf{H} = (1/\mu\omega)\mathbf{k}\times\mathbf{E}$.

APPENDIX B: CLOSED-FORM ANALYTICAL EXPRESSIONS FOR LOCAL INTEGRALS OF THE LAYERED GREEN'S FUNCTION

As mentioned in Sec. IV B 2, condition (47) must be met to compute integral (44) in the form (46).

Case 1: If condition (47) is met, the integrals in Eq. (45) can be computed as follows [I_1 is as in Eq. (46)]:

$$I_u = w_0\Delta v \left[\left(\frac{u_0-u_1}{k_u\psi_1 2} + \frac{1}{k_u 2\psi_1}\right)\left(\frac{1-e^{i\phi_1}}{\phi_1} - \frac{1-e^{i\phi_0}}{\phi_0}\right) \right.$$
$$- \left(\frac{u_0-u_2}{k_u\psi_2 2} + \frac{1}{k_u 2\psi_2}\right)\left(\frac{1-e^{i\phi_2}}{\phi_2} - \frac{1-e^{i\phi_0}}{\phi_0}\right)$$
$$+ \frac{u_1}{k_u\psi_1}\left(\frac{1-e^{i\phi_1}}{\phi_1 2} + \frac{ie^{i\phi_1}}{\phi_1}\right)$$
$$- \frac{u_0[k_v(\Delta v)2 + k_u(u_0-u_1)2]}{k_u\psi_1 2}\left(\frac{1-e^{i\phi_0}}{\phi_0 2} + \frac{ie^{i\phi_0}}{\phi_0}\right)$$
$$- \frac{u_2}{k_u\psi_2}\left(\frac{1-e^{i\phi_2}}{\phi_2 2} + \frac{ie^{i\phi_2}}{\phi_2}\right)$$
$$\left. + \frac{u_0[k_v(\Delta v)2 + k_u(u_0-u_2)2]}{k_u\psi_2 2}\left(\frac{1-e^{i\phi_0}}{\phi_0 2} + \frac{ie^{i\phi_0}}{\phi_0}\right) \right], \tag{B1}$$

$$I_v = \frac{w_0\Delta v}{k_u}\left[\frac{\Delta v}{\psi_1 2}\left(\frac{1-e^{i\phi_1}}{\phi_1} - \frac{1-e^{i\phi_0}}{\phi_0}\right) \right.$$
$$- \frac{\Delta v}{\psi_2 2}\left(\frac{1-e^{i\phi_2}}{\phi_2} - \frac{1-e^{i\phi_0}}{\phi_0}\right) + \frac{v_1}{\psi_1}\left(\frac{1-e^{i\phi_1}}{\phi_1 2} + \frac{ie^{i\phi_1}}{\phi_1}\right)$$
$$- \frac{v_2}{\psi_1}\left(\frac{1-e^{i\phi_0}}{\phi_0 2} + \frac{ie^{i\phi_0}}{\phi_0}\right) - \frac{v_1}{\psi_2}\left(\frac{1-e^{i\phi_2}}{\phi_2 2} + \frac{ie^{i\phi_2}}{\phi_2}\right)$$
$$\left. + \frac{v_2}{\psi_2}\left(\frac{1-e^{i\phi_0}}{\phi_0 2} + \frac{ie^{i\phi_0}}{\phi_0}\right) \right], \tag{B2}$$

$$I_w = \frac{w_0 2\Delta v}{k_u}\left[\frac{1}{\psi_1}\left(\frac{1-e^{i\phi_1}}{\phi_1 2} + \frac{ie^{i\phi_1}}{\phi_1} - \frac{1-e^{i\phi_0}}{\phi_0 2} + \frac{ie^{i\phi_0}}{\phi_0}\right) \right.$$
$$\left. - \frac{1}{\psi_2}\left(\frac{1-e^{i\phi_2}}{\phi_2 2} + \frac{ie^{i\phi_2}}{\phi_2} - \frac{1-e^{i\phi_0}}{\phi_0 2} + \frac{ie^{i\phi_0}}{\phi_0}\right) \right]. \tag{B3}$$

If, however, some part of the condition (47) is violated, another form should be used.

Case 2: If, $k_u \to 0$, then $\psi_1 = \psi_2 = \phi_v$, $\phi_1 = \phi_2$, and the exponentials in the integrands simplify to $e^{i(k_v v + k_w w)}$.

If $\phi \neq 0 \, \forall \, \phi \in \Phi_1$, where $\Phi_1 = \{k_v, \phi_0, \phi_1\}$, then one needs to use the second set of closed-form expressions for I_1, I_u, I_v, I_w. Detailed formulas for specific conditions are derived in Ref. [59].

Case 3: $k_u \to 0$, $k_v \to 0$ (then $k_w \neq 0$, $\phi_1 = \phi_2 = \phi_0 = \phi_w \neq 0$).

Case 4: $k_u \to 0$, $k_w \to 0$, $v_1 \neq 0$, and $v_2 \neq 0$.

Case 5: $k_u \to 0$, $k_w \to 0$, $v_1 \to 0$.

Case 6: $k_u \to 0$, $k_w \to 0$, $v_2 \to 0$.

Case 7: $k_u \to 0$, $\phi_1 \to 0$, and $k_w \neq 0$ (then $k_v \neq 0$, $v_1 \neq 0$, $\phi_0 \neq 0$).

Case 8: $k_u \to 0$, $\phi_0 \to 0$, and $k_w \neq 0$ (then $k_v \neq 0$, $v_2 \neq 0$, $\phi_1 \neq 0$).

Even when $k_u \neq 0$, other quantities in Φ may violate the condition (47).

Case 9: If $\psi_1 = \phi_v + k_u(u_0-u_1) \to 0$ (then $\psi_2 \neq 0$), and $\chi_1 \neq 0$, $\chi_2 \neq 0$, where $\chi_1 = k_u u_1 - k_v v_1 + \phi_w$, $\chi_2 = k_u u_2 - k_v v_1 + \phi_w$.

Case 10: $\psi_1 \to 0$, $\chi_1 \to 0$ (then $\chi_2 \neq 0$).

Case 11: $\psi_1 \to 0$, $\chi_2 \to 0$ (then $\chi_1 \neq 0$).

When $\psi_2 \to 0$ the proper integrals are chosen for the above three possible cases with χ_1 and χ_2. When $\phi_0 \to 0$, $\phi_1 \to 0$, or $\phi_2 \to 0$ the integrals should also be treated similarly as special cases.

[1] R. Shelby, D. R. Smith, and S. Schultz, Science **292**, 77 (2001).

[2] D. R. Smith, S. Schultz, P. Markos, and C. M. Soukoulis, Phys. Rev. B **65**, 195104 (2002).

[3] E. Kretschmann and H. Raether, Z. Naturforsch. A **23A**, 2135 (1968).

[4] A. Otto, Z. Phys. **216**, 398410 (1968).

[5] V. M. Agranovich and D. L. Mills, *Surface Polaritons: Electromagnetic Waves at Surfaces and Interfaces*, Modern Problems in Condensed Matter Sciences (North-Holland, Amster-

dam, 1982).

[6] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, Springer Tracts in Modern Physics No. 111 (Springer-Verlag, Berlin, 1988).

[7] A. Liebsch, *Electronic Excitations at Metal Surfaces* (Plenum Press, New York, 1997).

[8] D. W. Pohl, in *Near-Field Optics and Surface Plasmon Polaritons*, Vol. 81 of Topics in Applied Physics, edited by S. Kawata (Springer-Verlag, Heidelberg, 2001), Chapter 1, pp. 1-13.

[9] T. Okamoto, in *Near-Field Optics and Surface Plasmon Po-

laritons, (Ref. [8]) Chap. 3, p. 97123.

[10] P. B. Johnson and R. W. Christy, Phys. Rev. B **6**, 4370 (1972).

[11] H. Azizbekyan, M. J. Jory, A. P. Hibbins, and J. R. Shambles, J. Mod. Opt. **47**, 1227 (2000).

[12] J. Homola, S. S. Yee, and G. Gauglitz, Sens. Actuators B **54**, 315 (1999).

[13] B. Hecht, H. Bielefeldt, L. Novotny, Y. Inouye, and D. W. Pohl, Phys. Rev. Lett. **77**, 1889 (1996).

[14] I. I. Smolyaninov, D. L. Mazzoni, and C. C. Davis, Phys. Rev. Lett. **77**, 3877 (1996).

[15] S. I. Bozhevolnyi and V. Coello, Phys. Rev. B **58**, 10899 (1998).

[16] J. A. Sanchez-Gil, Appl. Phys. Lett. **73**, 3509 (1998).

[17] J. A. Sanchez-Gil and A. A. Maradudin, Phys. Rev. B **60**, 8359 (1999).

[18] R. Petit, *Electromagnetic Theory of Gratings* (Springer-Verlag, New York, 1980).

[19] E. Loewen and E. Popov, *Diffraction Gratings and Applications* (Marcel Dekker, New York, 1997).

[20] M. Nevière and E. Popov, *Light Propagation in Periodic Media: Differential Theory and Design* (Marcel Dekker, New York, 2002).

[21] E. Popov, B. Chernov, M. Neviere, and N. Bonod, J. Opt. Soc. Am. A **21**, 199 (2004).

[22] M. Moharam and T. Gaylord, J. Opt. Soc. Am. **72**, 1385 (1982).

[23] L. Li, J. Opt. Soc. Am. A **14**, 2758 (1997).

[24] C. Zhou and L. Li, J. Opt. A, Pure Appl. Opt. **6**, 43 (2004).

[25] L. Li, J. Chandezon, G. Granet, and J.-P. Plumey, Appl. Opt. **38**, 304 (1999).

[26] L. Li, J. Opt. Soc. Am. A **16**, 2521 (1999).

[27] A. Taflove and S. Hagness, *Computational Electrodynamics: The Finite Difference Time Domain Method* (Artech House, Boston, 2000).

[28] H. H. H. Steyskal, IEEE Trans. Antennas Propag. **47**, 1508 (1999).

[29] B. A. Munk, *Frequency Selective Surfaces* (John Wiley and Sons, New York, 2000).

[30] L. I. Goray, Proc. SPIE **4291**, 13 (2001).

[31] L. I. Goray, Proc. SPIE **5168**, 260 (2003).

[32] J. Seely, L. Goray, D. Windt, B. Kjornrattanawanich, Y. Uspenskii, and A. Vinogradov, Proc. SPIE **5538**, 43 (2004).

[33] L. I. Goray, Nucl. Instrum. Methods Phys. Res. A **536**, 211 (2005).

[34] B. Kleenmann and J. Erxmeyer, J. Mod. Opt. **51**, 2093 (2004).

[35] B. H. Kleenman, A. Mitreiter, and F. Wyrowski, J. Mod. Opt. **46**, 1323 (1996).

[36] J. P. Kottmann and O. J. F. Martin, IEEE Trans. Antennas Propag. **48**, 1719 (2000).

[37] J. P. Kottmann and O. J. F. Martin, Opt. Express **8**, 655 (2001).

[38] J. P. Kottmann, O. J. F. Martin, D. R. Smith, and S. Schultz, Phys. Rev. B **64**, 235402 (2001).

[39] J. P. Kottman and O. J. F. Martin, Appl. Phys. B: Lasers Opt. **74**, 299 (2001).

[40] C. Trampel, G. Kobidze, and B. Shanker, in *Proceedings of IEEE Symposium on Antenna and Propagation & URSI*, 2003, Vol. 30, p. 3159.

[41] K. A. Michalski and D. Zheng, IEEE Trans. Antennas Propag. **38**, 335 (1990).

[42] W. C. Chew, *Waves and Fields in Inhomogeneous Media* (Van Nostrand Reinhold, New York, 1990).

[43] J. A. Kong, *Electromagnetic Wave Theory* (John Wiley and Sons, New York, 1986).

[44] L. Felsen and N. Marcuvitz, *Radiation and Scattering by Waves* (Wiley-IEEE, New York, 1994).

[45] A. Ishimaru, *Electromagnetic Wave Propagation, Radiation, and Scattering* (Prentice-Hall, Englewood Cliffs, NJ, 1991).

[46] T. G. Livernois, D. P. Nyquist, and M. J. Cloud, IEEE Trans. Microwave Theory Tech. **39**, 579 (1991).

[47] D. P. Nyquist, J. M. Grimm, D. J. Infante, and H. Braunisch, Electromagnetics **17**, 105 (1997).

[48] D. H. Schaubert, D. R. Wilton, and A. W. Glisson, IEEE Trans. Antennas Propag. **32**, 77 (1984).

[49] G. Kobidze and B. Shanker, IEEE Trans. Antennas Propag. **52**, 2650 (2004).

[50] G. Kobidze, G. Jun, B. Shanker, and E. Michielssen, IEEE Trans. Antennas Propag. **53**, 1215 (2005).

[51] D. Wilton, S. M. Rao, A. Glisson, D. Schaubert, O. Al-Bundak, and C. Butler, IEEE Trans. Antennas Propag. **32**, 276 (1988).

[52] A. F. Peterson, S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*, IEEE/OUP Series on Electromagnetic Wave Theory (IEEE Press, New York, 1998), 1st ed.

[53] J. Humliček, J. Quant. Spectrosc. Radiat. Transf. **21**, 309 (1979).

[54] J. Humliček, J. Quant. Spectrosc. Radiat. Transf. **27**, 437 (1982).

[55] Z. Shippony and W. G. Read, J. Quant. Spectrosc. Radiat. Transf. **50**, 635 (1993).

[56] V. Lomakin and E. Michielssen, Phys. Rev. B **71**, 235117 (2005).

[57] A. G. Schuchinsky, D. E. Zelenchuk, and A. M. Lerer, J. Opt. A, Pure Appl. Opt. **7**, 102 (2005).

[58] H. F. Ghaemi, T. Thio, D. E. Grupp, T. W. Ebbesen, and H. J. Lezec, Phys. Rev. B **58**, 6779 (1998).

[59] G. Kobidze, B. Shanker, and D. P. Nyquist, Technical Report No. 2005-5, Michigan State University, 2005, (unpublished), http://www.egr.msu.edu/ece/