Fast high-order perturbation of surfaces methods for simulation of multilayer plasmonic devices and metamaterials

David P. Nicholls,* Fernando Reitich, Timothy W. Johnson, and Sang-Hyun Oh

1. INTRODUCTION

The scattering of time-harmonic linear waves by periodic media arises in a wide array of applications from materials science and nondestructive testing to remote sensing and oceanography. In this work we have in mind applications in optics, more specifically plasmonics, and the surface plasmon polaritons that are at the heart of remarkable phenomena such as extraordinary optical transmission, surface-enhanced Raman scattering, and surface plasmon resonance biosensing. In this paper we develop robust, highly accurate, and extremely rapid numerical solvers for approximating solutions to grating scattering problems in the frequency regime where these are commonly used. For piecewise-constant dielectric constants, which are commonplace in these applications, surface formulations are clearly advantaged as they posit unknowns supported solely at the material interfaces.

The algorithms we develop here are high-order perturbation methods and generalize previous approaches to take advantage of the fact that these algorithms can be significantly accelerated when some or all of the surfaces are trivial (flat). More specifically, for configurations with one nontrivial surface (and one trivial interface) we describe an algorithm that has the same computational complexity as a two-layer solver. With numerical simulations and comparisons with experimental data, we demonstrate the speed, accuracy, and applicability of our new algorithms. © 2014 Optical Society of America

Received April 4, 2014; revised June 20, 2014; accepted June 27, 2014; posted July 2, 2014 (Doc. ID 209329); published July 24, 2014

The scattering of time-harmonic linear waves by periodic media arises in a wide array of applications from materials science and nondestructive testing to remote sensing and oceanography. In this work we have in mind applications in optics, more specifically plasmonics, and the surface plasmon polaritons that are at the heart of remarkable phenomena such as extraordinary optical transmission (EOT) [7,8], surface-enhanced Raman scattering (SERS) [9], and surface plasmon resonance (SPR) biosensing [10].

A surface plasmon polariton (SPP) is usually defined as a time-harmonic electromagnetic wave propagating at the interface between a dielectric (insulator) and a metal (conductor), which is exponentially confined in the direction orthogonal to the interface. The configuration we study here is composed of a thin layer of gold (larger than the skin depth) mounted on (and illuminated through) a dielectric (e.g., glass or polymer substrate) with the exposed (flat) metal interface sitting in a second dielectric (e.g., water); see Fig. 1. These SPPs are the result of a resonance (a SPR) between the illuminating radiation and the polariton guided mode.

A remarkable feature of these SPRs is that for a well-designed (i.e., the metal should be a superior conductor, e.g., with high real negative permittivity) but fixed configuration, they are excited only for a very narrow band of illumination frequencies. For this reason it is easy to see how such devices could be constructed to produce measurement devices with very high sensitivity. Furthermore, the set of frequencies at which these SPRs exists depends strongly on the refractive index near the interface, and this principle is utilized in commercial SPR sensors that can detect refractive index changes that are as small as $10^{-7}$ refractive index units [10].

A well-known property of crucial importance is that these SPPs cannot be excited at the boundary of a dielectric/metal grating structure with flat interface [11,12]. It is not difficult to show that there is insufficient “momentum” to generate an SPP; however, with the addition of periodic corrugations at the interface such momentum can be provided and SPPs can be excited. (We note that periodic gratings are among several methods for producing SPPs, and we refer the interested reader to [11,12] for other approaches.)

With all of this in mind it becomes crucial to develop robust, highly accurate, and extremely rapid numerical solvers for approximating solutions to grating scattering problems in the resonance regime (both the period of the structure and the illumination frequency are in the optical regime of hundreds of nanometers) [13]. For piecewise-constant dielectric constants, which are commonplace in these applications, surface formulations are clearly advantaged as they posit unknowns supported solely at the material interfaces.
Methods based upon integral equation (IE) formulations are natural candidates \([14]\), but face several challenges. First, specially designed quadrature rules must be designed to deliver high-order (spectral) accuracy. Second, such rules generate dense, nonsymmetric-positive-definite linear systems to be solved. However, these issues have been adequately addressed (possibly with the use of iterative solution procedures accelerated by fast multipole methods), and they are a compelling alternative (see, e.g., the survey article of \([15]\) for more details). However, two properties render them noncompetitive for the periodic, parametrized problems we consider as compared with the methods we outline in this paper:

1. For periodic problems the relevant Green function must be periodized if one is to restrict the domain of integration to a single period cell. This is a well-known problem (see, e.g., the introduction of \([16]\) for a full description), and the slow convergence of the periodization must be accelerated (e.g., with techniques such as Ewald summation). However, even with such technology, these IE methods demand an additional discretization parameter: the number of terms retained in the approximation of the periodized Green function.

2. For configurations parametrized by the real value \(\epsilon\) (here the height/slope of the irregular interface), an IE solver will return the scattering returns only for a particular value of \(\epsilon\). If this value is changed, then the solver must be run again.

As we shall see, our new approach not only requires no special treatment for periodic problems, but also delivers the scattering returns for the entire family of configurations parameterized by the height/slope \(\epsilon\) (within the disk of analyticity of the relevant Taylor series) with a single simulation.

The alternative surface algorithm we have in mind is a high-order perturbation of surfaces (HOPS) method, which traces its roots to the low-order calculations of Rayleigh \([17]\) and Rice \([18]\). Its high-order incarnation [the method of variation of boundaries, later renamed the method of field expansions (FE)] for doubly layered media was first introduced by Bruno and Reitich \([19–21]\), and was further enhanced and stabilized by Nicholls and Reitich \([22–24]\). The method was expanded to multiple layers by Malcolm and Nicholls \([25]\) at the (reasonable) cost of requiring pairs of surface unknowns at every interface. Since these methods utilize the eigenfunctions of the Laplacian (suitable complex exponentials) on a periodic domain, the quasi-periodicity of solutions is “built in,” and, in contrast with IE methods, it does not need to be further approximated. Furthermore, since the methods are built upon expansions in the boundary parameter, \(\epsilon\), once the Taylor coefficients are known for the scattering quantities, it is simply a matter of summing these (rather than beginning a new simulation) for any given choice of \(\epsilon\) to recover the returns.

In the present contribution we extend the above contributions in a number of directions. First, we show that if one of the interfaces is flat (trivial), then the cost of the original two-layer (one irregular interface) algorithm can be recovered for the three-layer configurations that arise in this study of SPPs. Additionally, the algorithm is extended to accommodate the complex-valued permittivities that arise for metals. Following our previous developments on these methods \([22–23]\) we also develop a new, high-order, and provably stable HOPS method [the method of transformed field expansions (TFE)], which may be mandated for large and/or rough surface deviations. While this comes at a moderate additional cost (due to a small volumetric discretization), it is often crucial for the accurate simulation of configurations of interest \([26]\).

The paper is organized as follows: in Section 2 we describe the relevant models for simulating SPRs, in particular the time-harmonic Maxwell equations (Section 2.1) and a domain decomposition that greatly simplifies our recursions and enhances our numerical schemes (Section 2.2). We describe our methods of FE in Section 3, and TFE in Section 4. In Section 5 we present our numerical results with a special discussion of numerical implementation issues in Section 5.1. Evidence of the convergence of our schemes is presented in Section 5.2, while our numerical simulations of an experimental biosensing coupler are discussed in Section 5.3.

2. GOVERNING EQUATIONS

As discussed in the survey book of Raether \([27]\), the scattering of electromagnetic fields by metals can be effectively modeled with the classical Maxwell equations (i.e., quantum mechanical effects are negligible even for structures on the order of nanometers). However, an important consideration is the fact that the dielectric function \(\epsilon\) depends strongly on the wavelength of the radiation interacting with the structure.

A. Time-Harmonic Maxwell’s Equations

In Fig. 1 we display a cross section of the problem configuration we have in mind, a \(z\)-invariant layered structure of three materials. Occupying the domains above the plane \(y = \tilde{h}\) and below the graph \(y = g(x)\) are dielectrics with indices of refraction \(n_y \in \mathbb{R}\) and \(n_w \in \mathbb{R}\), while placed between is a (thin) layer of metal with refractive index \(n_e \in \mathbb{C}\):

\[
S := S_u \cup S_v \cup S_w
= \{y > \tilde{h}\} \cup \{g(x) < y < \tilde{h}\} \cup \{y < g(x)\}.
\]
The \( d \)-periodic grating, shaped by \( y = g(x) \), satisfies \( g(x + d) = g(x) \).

Incident upon this is monochromatic plane-wave radiation of frequency \( \omega \), aligned with the grooves of the grating and, in order to match with experimental results, impinging from negative infinity

\[
\mathcal{E}^i(x, y, t) = \mathbf{A}e^{j\omega t+ik_yy-i\beta_x x}, \quad \mathcal{H}^i(x, y, t) = \mathbf{B}e^{j\omega t+ik_yy-i\beta_x x}.
\]

Considering the reduced electric and magnetic fields

\[
\mathbf{E}(x, y) = \hat{\mathbf{e}} \mathcal{E}^i(x, y, t), \quad \mathbf{H}(x, y) = \hat{\mathbf{e}} \mathcal{H}^i(x, y, t),
\]

this \( \alpha \)-quasi-periodic illumination induces the same quasi-periodicity in the (reduced) scattered fields. Additionally, the scattered radiation must be “outgoing” (upward propagating in the upper layer \( S_u \) and downward propagating in the lower layer \( S_w \)), which, in this present context, is equivalent to solutions being bounded.

It is well known ([12], Chap. 2), that in this two-dimensional configuration, the time-harmonic Maxwell’s equations decouple into two scalar Helmholtz problems governing the transverse electric (TE) and transverse magnetic (TM) polarizations. These correspond to the components of the scattered electric and magnetic fields aligned with the invariant \( (z) \) directions, and we will denote them in the upper, middle, and lower layers by

\[
\mathbf{u} = \mathbf{u}(x, y), \quad \mathbf{v} = \mathbf{v}(x, y), \quad \mathbf{w} = \mathbf{w}(x, y),
\]

respectively. The incident radiation in the lowest layer is denoted by \( \mathbf{w}^i = \mathbf{w}^i(x, y) \).

These considerations lead us to consider \( \alpha \)-quasi-periodic, outgoing solutions of the system of boundary value problems

\[
\Delta \mathbf{u} + k_u^2 \mathbf{u} = 0 \quad y > \tilde{h}, \tag{1a}
\]

\[
\Delta \mathbf{v} + k_v^2 \mathbf{v} = 0 \quad g(x) < y < \tilde{h}, \tag{1b}
\]

\[
\mathbf{u} - \mathbf{v} = 0, \quad \partial_y \mathbf{u} - \tau^2 \partial_y \mathbf{v} = 0 \quad y = \tilde{h}, \tag{1c}
\]

\[
\Delta \mathbf{w} + k_w^2 \mathbf{w} = 0 \quad y < g(x), \tag{1d}
\]

\[
\mathbf{v} - \mathbf{w} = \zeta \quad y = g(x), \tag{1e}
\]

\[
\partial_N \mathbf{v} - \sigma^2 \partial_N \mathbf{w} = \psi \quad y = g(x), \tag{1f}
\]

where \( k_j = n_j \omega/c, N = (-\partial_x g, 1)^T \), the Dirichlet and Neumann data are

\[
\zeta(x) := \mathbf{w}^i(x, g(x)) = e^{j\omega g(x)} e^{j\alpha x}, \tag{1g}
\]

\[
\psi(x) := \sigma^2 (\partial_N w^i)(x, g(x)) = \sigma^2 (\partial_N g)(x) e^{j\beta g(x)} e^{j\alpha x}. \tag{1h}
\]

and the constants \( \tau^2 \) and \( \sigma^2 \) are identity in the TE configuration, and

\[
\tau^2 = \frac{k_u^2}{k_w^2} = \frac{n_u^2}{n_w^2}, \quad \sigma^2 = \frac{k_v^2}{k_w^2} = \frac{n_v^2}{n_w^2}, \tag{1i}
\]

in the TM case. It is a classical argument (see [27] for full details) that the case of TM polarization is the relevant one for the study of SPRs, and we restrict our attention to this from here.

**B. Domain Decomposition**

It is standard in the application of HOPS to affect a domain decomposition before expansion of field quantities in the boundary deformation. This can be achieved “exactly” with the use of Dirichlet–Neumann operators (DNOs) [22,28] in a “transparent boundary condition.” In Appendix A we give details of such conditions for the present configuration.

In summary, if we specify hyperplanes \( \{y = -b\} \) and \( \{y = a\} \), where

\[
-b < -|g|_{L^\infty}, \quad |g|_{L^\infty} < a < \tilde{h},
\]

then given DNOs \( T \) and \( S \), the system \( (1) \) is equivalent to

\[
\Delta \mathbf{v} + k_v^2 \mathbf{v} = 0 \quad g(x) < y < a, \tag{2a}
\]

\[
\partial_y \mathbf{v} - T[v] = 0 \quad y = a, \tag{2b}
\]

\[
\Delta \mathbf{w} + k_w^2 \mathbf{w} = 0 \quad -b < y < g(x), \tag{2c}
\]

\[
\mathbf{v} - \mathbf{w} = \zeta \quad y = g(x), \tag{2d}
\]

\[
\partial_N \mathbf{v} - \sigma^2 \partial_N \mathbf{w} = \psi \quad y = g(x), \tag{2e}
\]

\[
\partial_y \mathbf{w} - S[w] = 0 \quad y = -b. \tag{2f}
\]

We point out that while the uppermost field \( u \) has disappeared from the statement of the problem, it is still implicitly present through the operator \( T \), and can be readily recovered once \( v \) is known in \( \{g(x) < y < a\} \).

**3. METHOD OF FIELD EXPANSIONS**

The first HOPS method we consider for the approximate solution of Eq. \( (2) \) [equivalently Eq. \( (1) \)] is the method of FE [19–21,29]. This approach pursues the consequences of setting \( g(x) \equiv f(x) \) with the knowledge [23,28,30–32] that if \( f \) is sufficiently smooth (e.g., \( C^2 \), \( C^{1+\delta} \) Lipschitz), then the scattered fields \( \{v, w\} \) will depend analytically upon \( \varepsilon \) (sufficiently small) so that, e.g.,

\[
v = v(x, y; \varepsilon) = \sum_{n=0}^{\infty} v_n(x, y)e^{in}, \tag{3a}
\]

\[
w = w(x, y; \varepsilon) = \sum_{n=0}^{\infty} w_n(x, y)e^{in}. \tag{3b}
\]

Before proceeding we note that it has been further shown that the domain of analyticity includes a neighborhood of the
entire real axis [33] so that these expansions are valid for arbitrarily large real values of \( \varepsilon \) (up to physical obstruction, e.g., \( \varepsilon \sqrt{[\varepsilon]} < \hat{h} \)).

Upon insertion of the forms (3) into (2), differentiation \( n \) times with respect to \( \varepsilon \) followed by evaluating \( \varepsilon = 0 \) yields

\[
\Delta v_n + k_n^2 v_n = 0 \quad 0 < y < a, \tag{4a}
\]

\[
\partial_y v_n - T[v_n] = 0 \quad y = a, \tag{4b}
\]

\[
\Delta w_n + k_n^2 w_n = 0 \quad -b < y < 0, \tag{4c}
\]

\[
v_n - w_n = \zeta_n - Q_n \quad y = 0, \tag{4d}
\]

\[
\partial_y v_n - \sigma^2 \partial_y w_n = \psi_n - R_n \quad y = 0, \tag{4e}
\]

\[
\partial_y w_n - S[w_n] = 0 \quad y = -b, \tag{4f}
\]

where

\[
\zeta_n = F_n(i\beta_n)^n e^{i\alpha x}, \tag{4g}
\]

\[
\psi_n = \sigma^2 F_{n+1}(i\beta_n)^{n+1} + (\partial_x f) F_{n-1}(i\beta_n)^{n-1} e^{i\alpha x}, \tag{4h}
\]

\[
F_n(x) := (f(x))^n / n!, \text{ and } Q_n = \sum_{m=0}^{n-1} F_{n-m}(\partial_y^{-m} u_m - \partial_y^{-m} v_m), \tag{4i}
\]

\[
R_n = \sum_{m=0}^{n-1} F_{n-m}(\partial_y^{-m+1} u_m - \sigma^2 \partial_y^{-m+1} v_m), \tag{4j}
\]

\[
- \sum_{m=0}^{n-1} (\partial_x f) F_{n-1-m}(\partial_y^{-m-1} u_m - \sigma^2 \partial_y^{-m-1} v_m). \tag{4j}
\]

Appealing to Rayleigh’s expansions, we note that \( \alpha \)-quasi-periodic solutions of Eqs. (4a) and (4c) are

\[
v_n(x, y) = \sum_{p=-\infty}^{\infty} \xi_{n,p} e^{i\beta_n y} + D_j e^{-i\beta_n y} e^{i\alpha x}, \tag{5a}
\]

\[
w_n(x, y) = \sum_{p=-\infty}^{\infty} \hat{\mu}_{n,p} e^{-i\beta_n y} + E_j e^{i\beta_n y} e^{i\alpha x}, \tag{5b}
\]

where

\[
\alpha_p := \alpha + (2\pi/d)p, \quad \beta_{jp} := \sqrt{k_j^2 - \alpha_p^2}, \quad p \in \mathcal{U}^{(j)}, \tag{5c}
\]

\[
i \sqrt{\alpha_p^2 - k_j^2}, \quad p \in \mathcal{U}^{(j)}.
\]

and the propagating modes are specified by

\[
\mathcal{U}^{(j)} = \{ p \in \mathbb{Z} | \alpha_p^2 < k_j^2 \}.
\]

In light of the fact that \( S = (-i\beta_{e,D}) \) (see Section 2A), it is easy to see from Eq. (4f) that \( E_n \equiv 0 \) so that, for clarity, we write \( \hat{\mu}_{n,p} = \hat{w}_{n,p} \). The analysis at \( y = a \) is more involved, but it is not difficult to show that Eq. (4d) demands that

\[
D_p = e^{i\beta_{e,p}}(i\beta_{e,p} - \hat{T}_p)/(i\beta_{e,p} + \hat{T}_p).
\]

We note that in the case in which the same material fills the top two layers (i.e., \( k_n = k_o \)), \( \hat{T}_p = (i\beta_{e,p}) \) and \( D_p = 0 \) so that the FE recursions of Bruno and Reitich [19] are recovered. Finally, Eqs. (4d) and (4e) give

\[
(1 + D_p) \hat{\xi}_{n,p} - \hat{w}_{n,p} = \zeta_{n,p} - \hat{Q}_{n,p}, \tag{6a}
\]

\[
(i\beta_{e,p}) (1 - D_p) \hat{\xi}_{n,p} - \sigma^2 (i\beta_{e,p}) \hat{w}_{n,p} = \psi_{n,p} - \hat{R}_{n,p}, \tag{6b}
\]

a system of two linear equations (at every \( p \)) that is uniquely solvable provided

\[
\sigma^2 (i\beta_{e,p}) (1 + D_p) - (i\beta_{e,p})(1 - D_p) \neq 0.
\]

### 4. METHOD OF TRANSFORMED FIELD EXPANSIONS

The second HOPS method we describe, the method of TFE, follows a slightly different philosophy, which delivers not only computational stability, but also provably convergent recursions, at the cost of slightly elevated computational complexity [23,28]. In this TFE approach two preliminary changes of variables are affected:

\[
x' = x, \quad y' = a \left( \frac{y - g(x)}{a - g(x)} \right), \quad g(x) < y < a,
\]

\[
x'' = x', \quad y'' = b \left( \frac{y - g(x)}{b + g(x)} \right), \quad -b < y < g(x),
\]

which map \( \{ g(x) < y < a \} \) to \( \{ 0 < y' < a \} \) and \( \{ -b < y < g(x) \} \) to \( \{ -b < y'' < 0 \} \), respectively. This domain-flattening change of coordinates is known as the C method [34] in electromagnetics and \( \sigma \) coordinates [35] in the atmospheric sciences. The inverse transform is easily seen to be

\[
x = x', \quad y = y' + g(x') \left( \frac{a - y'}{a} \right), \quad 0 < y' < a,
\]

\[
x = x'', \quad y = y'' + g(x'') \left( \frac{b + y''}{b} \right), \quad -b < y'' < 0,
\]

which allows us to define

\[
v'(x', y') := v(x', y' + g(a - y')/a),
\]

\[
w'(x'', y'') := w(x'', y'' + g(b + y'')/b).
\]

For ease of exposition, we will, from this point forward, drop reference to the primed variables.
In these new coordinates we find that Eq. (2) is transformed to
\[
\Delta v + k_v^2 v = F^v(x, y; g, v) \quad 0 < y < a, \tag{7a}
\]
\[
\partial_y v - T[v] = J^v(x; g, v) \quad y = a, \tag{7b}
\]
\[
\Delta w + k_w^2 w = F^w(x, y; g, w) \quad -b < y < 0, \tag{7c}
\]
\[
v - w = \zeta - Q(x; g, v, w) \quad y = 0, \tag{7d}
\]
\[
\partial_y v - \sigma^2 \partial_y w = \psi - R(x; g, v, w) \quad y = 0, \tag{7e}
\]
\[
\partial_y w - S[w] = J^w(x; g, w) \quad y = -b, \tag{7f}
\]

The particular forms for $F^v$, $J^v$, $F^w$, $Q$, $R$, and $J^w$ have been derived in the work on two-layer configurations [36]. The formula for $J^v$ is necessarily new as we consider a three-layer structure here; however, it is not significantly different from the term found in [36], and we present it without derivation:
\[
J^v = -\frac{1}{a} \partial^2 T[v].
\]

The next step in the TFE procedure is, upon defining $g(x) = ef(x)$, to expand the transformed fields $\{v, w\}$ in the Taylor series
\[
v(x, y; e) = \sum_{n=0}^{\infty} v_n(x, y) e^n, \quad w(x, y; e) = \sum_{n=0}^{\infty} w_n(x, y) e^n,
\]
which can be shown to be strongly convergent in an appropriate Sobolev space [28, 33]. Upon insertion of these forms into Eq. (2) we find, at each perturbation order $n$, the following problem to solve:
\[
\Delta v_n + k_v^2 v_n = F^v_n(x, y) \quad 0 < y < a, \tag{9a}
\]
\[
\partial_y v_n - T[v_n] = J^v_n(x) \quad y = a, \tag{9b}
\]
\[
\Delta w_n + k_w^2 w_n = F^w_n(x, y) \quad -b < y < 0, \tag{9c}
\]
\[
v_n - w_n = \zeta_n - Q_n \quad y = 0, \tag{9d}
\]
\[
\partial_y v_n - \sigma^2 \partial_y w_n = \psi_n - R_n(x) \quad y = 0, \tag{9e}
\]
\[
\partial_y w_n - S[w_n] = J^w_n(x) \quad y = -b, \tag{9f}
\]

It is not difficult (cf. [36]) to derive the forms for $\{F^v_n, J^v_n, F^w_n, Q_n, R_n, J^w_n\}$, which we exclude for brevity.

5. NUMERICAL RESULTS

In this section we present a brief sampling of the results our new algorithms can achieve. We begin with a configuration of three insulators (dielectrics) that do not support SPRs, but that are useful as they possess a principle of conservation of energy that is a standard diagnostic of convergence. Having validated our codes with this initial configuration of insulators, we proceed to a structure that does exhibit SPRs: a triply layered water/gold/dielectric configuration that we further compare with experimental observations.

Before beginning, we describe the “outputs” of our simulations, which are quantities of particular interest to experimentalists. The notion of a “far-field pattern” is pervasive in the study of scattering by bounded obstacles [37] but is not really relevant for gratings. The quantities that play the same role in these configurations are the efficiencies. For the problem we describe in Eq. (1) they are defined in terms of the fields in the bottom (c) and top (u) layers. To be more specific, the Rayleigh expansions state that
\[
u(x, y) = \sum_{p=-\infty}^{\infty} \tilde{u}_p e^{i p y} e^{i k_p x}, \quad y > \tilde{h},
\]
\[
w(x, y) = \sum_{p=-\infty}^{\infty} \tilde{w}_p e^{i p y} e^{i k_p x}, \quad y < \tilde{b} < |-y|_{L^2},
\]
and in terms of these we can define the efficiencies, for illumination from below the structure, as
\[
e_{u,p} = \frac{\tilde{u}_p}{\tilde{w}_p} |\tilde{u}_p|^2, \quad e_{w,p} = \frac{\tilde{w}_p}{\tilde{w}_p} |\tilde{w}_p|^2.
\]

These efficiencies are the (scaled) amplitudes of the transmitted and reflected waves at the propagating frequencies $p$ in $U^{(v)}$ and $U^{(w)}$, and, thus, they describe the information “seen in the far field.” If all three materials in the structure are lossless (i.e., perfect insulators), then there is a conserved energy
\[
\sum_{p \in L^2(v)} e_{u,p} + \frac{k_v^2}{k_w^2} \sum_{p \in L^2(w)} e_{w,p} = 1,
\]
and we can define a diagnostic of convergence based upon this, namely the “energy defect”:
\[
\delta := 1 - \sum_{p \in L^2(v)} e_{u,p} - \frac{k_v^2}{k_w^2} \sum_{p \in L^2(w)} e_{w,p}. \tag{10}
\]

Regardless of the material’s index of refraction, it is of considerable interest to know how much energy is reflected and/or transmitted by a structure. Since the direction of illumination of the structure may be from above or below, we do not distinguish between reflection or transmission, but rather define the “reflectivity” of the scattered field in the upper and lower layers. These are defined as
\[
\tilde{R}^{u} := \sum_{p \in L^2(v)} e_{u,p}, \quad \tilde{R}^{w} := \sum_{p \in L^2(w)} e_{w,p},
\]
so that for a stack of insulators $\tilde{R}^{u} + (k_v^2/k_w^2)\tilde{R}^{w} = 1$. Of great importance are the “reflectivity maps,” which are these
quantities as functions of the wavelength of the incident radiation, \( \lambda \), and the size (height/slope) of the interface deformation, \( h \). It is often of interest to compare these reflectivity maps for nontrivial interfaces (of size \( h \)) with those realized by the same structure with flat interfaces; thus we define the “scaled reflectivity maps,” e.g.,

\[
R^w = R^w(\lambda, h) := \frac{R^w(\lambda, h)}{R^w(\lambda, 0)}.
\] (11)

Finally, two scattered directions that are of special interest to practitioners are the specular ones (the directions of reflection and transmission in the flat-interface case). In terms of our representations, these energies are given by \( e_{u,0} \) and \( e_{w,0} \), which, in light of the formulas above, are given by \(|\hat{u}_0|^2\) and \(|\hat{w}_0|^2\). Again, of particular import are the “scaled specular energies,” e.g.,

\[
C_0 = C_0(\lambda, h) := \frac{|\hat{u}_0(\lambda, h)|^2}{|\hat{w}_0(\lambda, 0)|^2}.
\] (12)

A. Numerical Implementation

For a numerical implementation of the FE method we approximate \( \{v, w\} \) by a truncation of the expansions (3)

\[
v \approx v^N := \sum_{n=0}^{N} v_n(x, y)e^{in\alpha}, \quad w \approx w^N := \sum_{n=0}^{N} w_n(x, y)e^{in\alpha}.
\] (13)

For the \( \{v_n, w_n\} \) we approximate by the \( N_n \)-term truncation of Eq. (5):

\[
v_n \approx v^N_n := \sum_{p=-N_n/2}^{N_n/2-1} \hat{v}_{n,p}e^{ip\beta x} + D_pe^{i\phi_0}e^{ip\beta y}e^{ip\xi x},
\] (14a)

\[
w_n \approx w^N_n := \sum_{p=-N_n/2}^{N_n/2-1} \hat{w}_{n,p}e^{-ip\beta y}e^{ip\phi x}.
\] (14b)

Finally, the \( \{\hat{v}_{n,p}, \hat{w}_{n,p}\} \) are recovered by solving (6) where the only approximation is that convolutions arising in the formulas for \( \{Q_{n,p}, R_{n,p}\} \) are evaluated by the discrete Fourier transform (DFT) accelerated by the fast Fourier transform (FFT) algorithm [38].

The implementation of the TFE algorithm is a little more involved, but begins analogously to the FE procedure (recalling that we dropped the primes in the TFE change of variables) by approximating

\[
v \approx v^N := \sum_{n=0}^{N} v_n(x, y)e^{in\alpha}, \quad w \approx w^N := \sum_{n=0}^{N} w_n(x, y)e^{in\alpha},
\]

and

\[
v_n \approx v^N_n := \sum_{p=-N_n/2}^{N_n/2-1} \hat{v}_{n,p}(y)e^{ip\beta x},
\]

Upon insertion of these forms into Eq. (9), it becomes apparent that we must solve a pair of coupled two-point boundary value problems for \( \{\hat{v}_{n,p}(y), \hat{w}_{n,p}(y)\} \), on the domain \([-b, 0] \cup [0, a]\) [36]. For this we diverge from the approach of [36] and utilize a Chebyshev-tau method outlined in [39]. In short, we express

\[
\hat{v}_{n,p} \approx \hat{v}^N_{n,p} := \sum_{l=0}^{N_n} \hat{v}_{n,p,l}T_l\left(\frac{2y - a}{a}\right),
\]

\[
\hat{w}_{n,p} \approx \hat{w}^N_{n,p} := \sum_{l=0}^{N_n} \hat{w}_{n,p,l}T_l\left(\frac{2y + b}{b}\right),
\]

and find the \( \{\hat{v}_{n,p,l}, \hat{w}_{n,p,l}\} \) from the Chebyshev-tau constraints. We point out that while this solution procedure can be accelerated by the FFT algorithm (the computational complexity is \( O(N_n \log(N_n)) \), it is disadvantaged when compared with the FE algorithm as an extra discretization is required in the \( y \) variable.

To conclude our discussion of numerical implementation we point out that there is a choice in evaluating the truncated Taylor series that appear above [see, e.g., Eq. (13)]. The classical numerical analytic continuation technique of Padé approximation [40] has been successfully brought to bear upon HOPS methods in the past (see, e.g., [20,33]), and we utilize this here as well. This approximant has the remarkable properties that, for a wide class of functions, not only is the convergence faster at points of analyticity, but it also may converge for points outside the disk of analyticity. We refer the interested reader to Section 2.8 of Baker and Graves-Morris [40] and the insightful calculations of Section 8.3 of Bender and Orszag [41] for a thorough discussion of the capabilities and limitations of Padé approximants.

B. Convergence Test

In this section we provide evidence for the accuracy and robustness of the FE and TFE methods we outlined in Sections 3 and 4. As there are no readily available exact solutions for plane-wave scattering by corrugated gratings we resort to two widely accepted measures of convergence: Cauchy convergence and energy defect.

Before describing these measures of convergence we outline the physical and numerical parameters of our simulations. In light of our subsequent experiments we choose the periodicity of our grating to be \( d = 650 \text{ nm} \) (0.65 \( \mu \text{m} \)), and the mean interface locations to be \( h = 1000 \text{ nm} \) and \( g = 0 \text{ nm} \). We fill this triply layered structure with three (different) perfect insulators with indices of refraction

\[
n_u = 1.1, \quad n_v = 2.1, \quad n_w = 3.5.
\]

For the shape of the lower interface we choose the “rough profile” selected in [33]
\[
f(x) = a \left\{ \cos \left( \frac{2\pi x}{d} \right) + \frac{1}{5} \cos \left( \frac{6\pi x}{d} \right) \\
+ \frac{1}{16} \cos \left( \frac{8\pi x}{d} \right) + \frac{1}{5} \sin \left( \frac{6\pi x}{d} \right) \right\},
\]

where \( a = 25 \text{ nm} \), and we chose 101 equally spaced values of \( h = \epsilon \) between 0 and 1. For the incident radiation we chose normal incidence (\( \alpha = 0 \)), and 101 equally spaced wavelengths, \( \lambda \), between 600 and 750 nm (0.6 and 0.75 \( \mu \text{m} \)). For numerical parameters we selected \( N_x = 128 \), \( N_y = 32 \), and \( N_{\text{max}} = 30 \).

For the FE algorithm we present in Fig. 2(a) measurements of the difference between the reflectivity map \( R^w \) with \( N \) Taylor terms versus this map with \((N - 1)\) Taylor terms measured in the supremum norm. We repeat this calculation for the TFE approach and display results in Fig. 2(b). In both instances we see the spectral convergence rate one would expect of the Fourier/Chebyshev/Taylor approach we outline above [33].

C. Buried Plasmonic Grating

We now consider the configuration that has been studied experimentally by two of the authors (Oh and Johnson) in [42] to investigate the thin-film sensing capability of an engineered metal film that is flat on one side but has a periodic grating patterned on the opposite side. (This is in contrast to conventional gratings in which undulations appear on both sides.) For this, a structure was fabricated consisting of a thin layer of gold mounted on a polymeric substrate (optical adhesive, through which the structure is illuminated) and sitting in water. The key feature of this structure is that surface reactions and molecular binding events occur on a flat surface.
whereas a grating coupler for SPP excitation is buried in the film and does not disturb the surface topography. There are broader applications of such buried grating structures; for example, a multilayer metamaterials stack can be placed on such buried gratings or slits [43], which can be readily modeled using our method. For this study, we focus on a single-layer metal film atop a buried grating, as illustrated in Fig. 1, without losing generality.

To accommodate this geometry we consider the upper (u) layer as water, the middle layer (v) as gold, and the lowest layer (w) as polymer (Norland 61). The configuration is depicted to scale in Fig. 4(a), and rescaled in Fig. 4(b) to reveal the features of the gold/substrate interface. The technology of [42] involves depositing a (thin) \( t_1 \) nm layer of gold, followed by \( t_2 \) nm strips of gold of width \( w \) nm, and finishing with an (effectively) infinite layer of epoxy. To approximate this we set \( h = (t_1 + t_2) \) nm, and shape the gold/epoxy interface by the function

\[
\theta(x) = a \tanh(b(x - d/2 + c)) - \tanh(b((x - d/2) - c)),
\]

where we have set \( a = t_2/2 \), \( c = (d - w)/2 \), and the dimensionless “steepness” \( b = 5 \times 10^5 \).

An important consideration is the model of the refractive index for each of these three layers. For the two dielectrics we use [42]

\[
\varepsilon_{\text{epoxy}} = 1.56, \quad \varepsilon_{\text{water}} = 1.333.
\]

The refractive index of gold is the subject of current research, and we choose a Lorentz model [44]

\[
\varepsilon_{\text{Au}} = \varepsilon_{\text{Au}}^{\infty} + \sum_{j=1}^{6} \frac{\Delta \varepsilon_{\text{Au}}}{\omega_j^2 - \omega^2 - i\gamma_{\text{Au}} \omega},
\]

where \( \omega = 2\pi/\lambda \), \( \varepsilon_{\text{Au}}^{\infty} = 1 \), and \( \Delta \varepsilon_{\text{Au}}, \gamma_{\text{Au}}, b_{\text{Au}}, \) and \( c_{\text{Au}} \) can be found in [44].

To demonstrate the applicability of our new algorithm we now make direct comparisons the experimental results presented in [42]. In Fig. 5 we plot the scaled specular energy \( C_0 \) [cf. Eq. (12)], which is reflected back into the epoxy. This figure includes not only experimental data (with green diamonds), but also the results of numerical simulations (in red crosses) using the rapid and robust FE algorithms over a range of incident wavelengths from \( \lambda = 550 \) to 750 nm. We find the agreement between the two curves quite striking, especially in light of the fact that these SEMs simply give estimates of these quantities. However, we do point out that while the SEM utilized after the experiment indicated that \( t_1 = 50 \) nm, \( t_2 = 20 \) nm, and \( w = 220 \) nm, we found that if we set \( t_1 = 40 \) nm, \( t_2 = 22.2 \) nm, and \( w = 209 \) nm we got the “best” results depicted here.

We note that there is a pronounced “dip” in the neighborhood of \( \lambda = 631 \) nm and a “peak” near \( \lambda = 717 \) nm, which was verified by comparison with the stable and high-order (though more computationally intensive) TFE recursions. For this reason we plot in Fig. 6 the intensity of the reflected field for

![Plot of Configuration](image)

**Fig. 4.** (a) Geometry and model (to scale). Upper region, water (sensing area); middle, gold; lower, polymer substrate. Gold/polymer interface specified by Eq. (15): blue curve displays \( b = \infty \), while green curve is “idealized” shape with \( b = \infty \). (b) Close up of (a) near the gold/polymer interface.

![Plot of Configuration](image)

**Fig. 5.** Plot of the reflectivity map for water/gold/polymer configuration (which can exceed 1 as it is normalized by the flat-interface value). Experimental data are depicted with green diamonds, while numerical simulations (via FE recursions) are shown with red stars.
Using the TFE algorithm (which delivers the scattered field everywhere in the problem domain, including the selvage region),

To close we report on two numerical experiments that probe the sensitivity of these devices. First, we repeat one of the simulations reported in [42] by plotting, in Fig. 7, the shift of the dip and peak data (from the value shown in Fig. 5) as the index of refraction, $n_{\text{water}}$, is varied from 1.33 to 1.36, which measures the bulk sensitivity and is achieved experimentally by varying the concentration of glycerol in the water.

Based upon a least squares fit to this data we find that the dip changes by 82.5 nm/RIU (“refractive index unit”), compared with roughly 300 nm/RIU reported in [42], while the peak changes by 491 nm/RIU compared with approximately 410 nm/RIU in [42].

Finally, we consider another simulation described in [42] by plotting, in Fig. 8, the spectral shift of the SPR resonances as a dielectric film is sequentially deposited layer by layer on top of the sensing surface. Such experiments are often performed, e.g., using atomic layer deposition (ALD) of dielectric films on metals [45], to measure the response of optical sensors as a function of the deposited film thickness. Because the intensity of the SPP evanescent field decreases rapidly away from the interface, the response of these sensors (e.g., the resonance wavelength) does not scale linearly with the deposited film thickness. To precisely model such effects, one needs an accurate modeling capability to resolve the changes due to a very thin (a few nanometers) film atop the patterned sensing surface. We point out that this requires a significant but straightforward extension of the formulation we have described above in that the boundary operator, $T$, must be generalized to the case of two flat interfaces (four layers total). This, and the further extension to arbitrary numbers of flat interfaces (in both two and three dimensions), is the subject of a forthcoming publication. Returning to our results, based upon a least squares fit to the first half of the data, we find that, in the range of 0–50 nm, the dip changes by 0.640 nm per nm of Al$_2$O$_3$ overlayer, while the peak varies by...
Neumann data can be computed as
\[ \partial_y \psi(x, -b) = \sum_{p=-\infty}^{\infty} (-i\beta_{p,p}) \psi_p e^{i\beta_p y} = : S[\psi(x)], \]
which defines the order-one Fourier multiplier \( S \) (commonly denoted \((-i\beta_{p,p})\)) that permits only downward propagating solutions. We point out for later reference that \( S \) is a DNO, and the Neumann condition in \((A1c)\) can be stated at \( y = -b \) in terms of this DNO as
\[ \partial_y \psi - S[\psi] = 0. \]
Thus Eq. \((A1)\) is equivalent to
\[ \Delta w + k_w^2 w = 0 \quad -b < y < g(x), \quad (A3a) \]
\[ \partial_y w - S[w] = 0 \quad y = -b. \quad (A3b) \]

The considerations for the transparent boundary condition above the structure are a little more involved as we locate the “artificial boundary” between the irregular interface and the uppermost interface
\[ |g|_{L^\infty} < a < \tilde{h}. \]
We now focus on the unknowns \( \{u, v\} \) in Eq. \((1)\) and quasi-periodic and outgoing solutions of the augmented system
\[ \Delta u + k_u^2 u = 0 \quad y > \tilde{h}, \quad (A4a) \]
\[ \Delta v + k_v^2 v = 0 \quad a < y < \tilde{h}, \quad (A4b) \]
\[ u - v = 0, \quad \partial_y (u - r^2 v) = 0 \quad y = \tilde{h}, \quad (A4c) \]
\[ \Delta v + k_v^2 v = 0 \quad g(x) < y < a, \quad (A4d) \]
\[ v - v = 0, \quad \partial_y (v - v) = 0 \quad y = a. \quad (A4e) \]

We now pursue a formulation of the problem for \( \{u, v\} \) in terms of surface Dirichlet
\[ U(x) := u(x, \tilde{h}), \quad V(x) := v(x, \tilde{h}), \quad V^a(x) := v(x, a), \]
and external Neumann traces
\[ \tilde{U}(x) := -(\partial_y u)(x, \tilde{h}), \quad \tilde{V}(x) := (\partial_y v)(x, \tilde{h}), \]
\[ \tilde{V}^a(x) := (\partial_y v)(x, a); \]
we point the interested reader to \([49]\) for further details. In terms of these Eqs. \((A4a)-(A4c)\) are equivalent to
\[ U - V = 0, \quad -\tilde{U} - r^2 \tilde{V} = 0, \quad V^a = \psi. \] (A5)

where we view \( \psi := \nu(x, a) \) as given and seek to produce \( (\partial_x \nu)(x, a) \), precisely the DNO \( T \). Further defining the DNOs \( (A9) \)

\[ G: U \to \tilde{U}, \quad B: (V, V^a) \to (\tilde{V}, \tilde{V}^a), \]

where we identify \( B \) as the matrix-valued operator

\[ B = \begin{pmatrix} B^{uu} & B^{ud} \\ B^{du} & B^{dd} \end{pmatrix}, \]

Eq. (A5) is equivalent to

\[ U - V = 0, \quad -G[U] - r^2(B^{uu}[V] + B^{ud}[V^a]) = 0, \quad V^a = \psi, \]

or, quite simply,

\[ (G + r^2B^{uu})[V] = -r^2B^{ud}[\psi]. \] (A6)

Now, the operator \( T \) is given by

\[ T[\psi] = -(B^{uu}[V] + B^{ud}[\psi]) = (r^2(G + r^2B^{uu})^{-1}B^{ud} - B^{dd})[\psi], \]

which, despite its intimidating form, can be expressed as a simple multiplication in Fourier space. This is because it has been shown that for the flat-interface case described here, the operators \( G \) and \( B \) are order-one Fourier multipliers with formulas

\[ G = -(i\beta_{\nu,D}), \quad B^{uu} = B^{dd} = (i\beta_{\nu,D}) \coth(i\beta_{\nu,D}(\bar{h} - a)), \]

\[ B^{ud} = B^{du} = -(i\beta_{\nu,D}) \text{csch}(i\beta_{\nu,D}(\bar{h} - a)), \]

so that

\[ \tilde{T}_p = -\frac{r^2(i\beta_{\nu,p}) \text{csch}(i\beta_{\nu,p}(\bar{h} - a))}{-(i\beta_{\nu,p}) + r^2(i\beta_{\nu,p}) \coth(i\beta_{\nu,p}(\bar{h} - a))} - (i\beta_{\nu,p}) \coth(i\beta_{\nu,p}(\bar{h} - a)). \] (A7)

Finally, we note that Eq. (A4) is equivalent to

\[ \Delta \nu + k_0^2 \nu = 0 \quad g(x) < y < a, \] (A8a)

\[ \partial_y \nu - T[\nu] = 0 \quad y = a. \] (A8b)

ACKNOWLEDGMENTS

D. P. N. gratefully acknowledges support from the National Science Foundation through grant no. DMS-1115333, and the Department of Energy under Award No. DE-SC0001549. F. R., T. W. J, and S.-H. O. were supported in part by the National Science Foundation through grant no. DMR-0941537. T. W. J and S.-H. O. also acknowledge support from the National Science Foundation through grant no. CBET-1067681. This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

REFERENCES


27. H. Raether, Surface Plasmons on Smooth and Rough Surfaces and on Gratings (Springer, 1988).