Oulton, Rupert F. and Pile, David F. P. and Liu, Yongmin and Zhang, Xiang (2007) Scattering of Surface Plasmon Polaritons at Abrupt Surface Discontinuities: Implications for Nano-Scale Cavities. *Physical Review B : Condensed Matter and Materials Physics* 76:035408.

Scattering of Surface Plasmon Polaritons at Abrupt Surface Discontinuities: Implications for Nano-Scale Cavities.

R. F. Oulton, D. F. P. Pile, Y. Liu and X. Zhang

NSF Nano-scale Science and Engineering Center,

*3112 Etcheverry Hall, University of California, Berkeley CA 94720.**

Abstract

We have developed a rigorous mode matching approach for the exact semi-analytical analysis of surface plasmon propagation across non-uniform semi-infinite dielectric-metal interfaces. We address two key deficiencies of related approaches in the literature: firstly, we resolve issues of accuracy and convergence and secondly, while we focus on the analysis of two-dimensional problems, we present a framework for three-dimensional problems for the first time. Analytical derivations of coupling coefficients between guided and radiation modes allow an efficient scattering matrix formulation to describe general structures with multiple discontinuities. Studies of the reflection, transmission and radiation of surface plasmons incident on both dielectric and metallic surface discontinuities show a correspondence with an effective Fresnel description. We also model a surface plasmon Distributed Bragg Reflector (DBR) capable of reflecting between 80 % and 90 % of incident surface plasmon power. Radiation mode scattering ultimately limits the DBR's reflection performance rather than the intrinsic absorption of the metal. Thus alternative plasmonic geometries that suppress radiation modes, such as gap and channel structures, could be superior for the design of strongly reflective DBRs for integration in high Q-factor nano-scale cavities. We anticipate that this method will be an invaluable tool for the efficient and intuitive design of plasmonic devices based on structural non-uniformities.

PACS numbers:

I. INTRODUCTION

Surface plasmon polaritons[1] have recently re-emerged as a unique and promising method for efficiently concentrating and delivering electromagnetic energy to the nano-scale: "plasmonics" now offers unique opportunities for sub-wavelength optical waveguides[2–4], new sensors and detection techniques based on surface-enhanced Raman scattering[5,6] and sub vacuum wavelength superlens imaging[7] amongst others. Most recently, Miyazaki and Kurokawa[8] experimentally demonstrated a nano-scale metallic gap-plasmon based cavity whose fundamental resonant length was a tenth of the vacuum wavelength. Such observations have re-invigorated new research into nano-scale light emitting devices such as surface plasmon amplification by stimulated emission of radiation[9] and nano-lasers.

Many authors have already experimentally and numerically examined components necessary to construct nano-scale cavities. For example, Krenn and co-workers introduced a series of 2 dimensional components based on periodic arrays of gold nano-particles for achieving extremely efficient Bragg reflection of surface plasmons[10]. Introducing surface discontinuities in this way naturally points to the potential to build surface wave cavities. An excellent review by Zayats discusses such aspects of surface wave optics[11]. A similar principle lies behind Miyazaki's approach[8]: here the discontinuities at the terminated ends of a gap plasmon waveguide provide the necessary modal reflection. Most recently, three numerical studies considered nano-scale thin-film[12,13] and channel plasmon waveguide[14] cavities employing some form of surface discontinuity to achieve cavity feedback. However, these past works are predominantly experimental or brute force studies that somewhat disguise important insight into these phenomena.

The present paper uses a rapidly converging and accurate semi-analytical mode matching[15] method for solving systems of plasmons interacting with multiple abrupt structural metal/dielectric half spaces transitions, such as those schematically illustrated in Fig. 1. Our mode matching approach demonstrates excellent convergence even for highly confined surface plasmon modes with significant penetration into the metal and directly provides modal coupling coefficients (*i.e.* reflection, R, transmission, T and scattering, S, $\bar{S}$). Two key deficiencies of existing semi-analytical approaches are dealt with: firstly, issues of accuracy and correct convergence and secondly, a framework for three dimensional problems (Sec. IV D). The current investigation tackles both these issues, but focuses mainly on the

first due to constraints of space. Strengths of surface wave reflection, R, transmission, T, and forward, S, and backward, $\bar{S}$, scattering at single and multiple abrupt discontinuities of dielectric (Sec. IV A) and metallic (Sec. IV B) permittivity are determined and analyzed. A surface plasmon Distributed Bragg Reflector (DBR) capable of reflecting between 80 % and 90 % of incident surface plasmon power is presented (Sec. IV C) and analyzed.

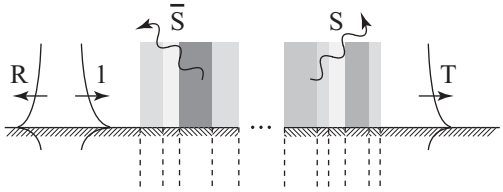

FIG. 1: Illustration of metallic / dielectric half space surface discontinuities.

Plasmonics research today employs a wide variety of numerical techniques for the analysis of highly confined electromagnetic fields. The most popular tools are the finite-difference time-domain (FDTD) and finite-element (FEM) methods. Unfortunately commercial implementations tend to be inflexible and don't clearly reveal the numerical methods employed and home-made versions require prodigious programming time and effort for any sophisticated analysis of complex structures. All FDTD and FEM algorithms require vast computational resources to solve even small three-dimensional (3D) problems. Moreover, these methods do not lend themselves well to intuitive analysis or clear interpretation of the resultant field profiles they generate. Analytical and semi-analytical methods are still applicable in many plasmonic systems; in these cases, the greater insight and computation efficiency afforded can be crucial for the development of novel devices. The most prevalent in current plasmonics literature are the Green functions approach[16,17] and Discrete Dipole Approximation approach[18,19]. While providing all the benefits of their analytical nature, these methods are extremely complicated to implement in general cases and are usually applied to the problem of scattering from single or multiple surface defects.

Schevenchko and co-workers (See Ref.[20] and references therein) originally laid out analytical methodology for describing open waveguide systems, such as that shown in Fig. 1, proving mathematical completeness of the normal mode expansion and describing approximate analytical solutions to problems involving continuous surface variations at Radio Frequencies (RF). Later, Mamoud and Beal[21] combined these analytical techniques with the numerical mode matching approach of Clarricoats and Slinn[15]. The motivation for their

work concerned the detection of abrupt discontinuities on dielectric loaded RF transmission lines, where an analytical solution was not tangible due to the highly coupled nature of the normal modes. Unfortunately, their result employed a very low number of radiation modes: with 2 or 3 Laguerre polynomials describing the radiation fields, the convergence of their approach is surprising in light of the numerical analysis in this paper. Stegeman *et al*[22] also solved essentially the same problem, but did not follow the mode matching approach and considered boundary matching the transverse fields at discrete points along a discontinuity. Their approach suffered two numerical issues: firstly, sufficient accuracy required a large number of boundary matching points, which consequently limited the number of normal modes that they could consider to four; secondly, in order to obtain convergence, the authors had to introduce an artificial grounding plate reducing the problem to essentially a closed waveguide system, which limits the description of scattering loss. The authors also commented on the poor convergence properties of treating the open waveguide system in a similar fashion to Mamoud and Beal, which, in light of the present study, certainly merits closer attention. Voronko and co-workers[23] considered the simpler problem of surface discontinuities of metal/dielectric half spaces with an approach similar to the mode matching method. However, the authors made two simplifications: firstly, they neglected inter-radiation mode scattering, which is not a valid approximation for large variations in permittivity at the discontinuity; and secondly, they solved the integral equations for the single interface directly, which limits further adaptation for the description of multiple discontinuities. Finally, as a general observation of the early work on this problem, detailed studies of specific results are missing from the literature and what studies do exist are for the RF band. The review by Zayats[11] summarizes the implementation of these methods and key results.

II. MODE MATCHING SURFACE WAVES.

The mode matching method[15] operates on the premise that a complete set of orthonormal modes on either side of the discontinuity under study is describable analytically. The completeness of the set ensures a consistent map between the mode expansions on either side of a discontinuity. Section II A derives the set of surface modes for use in the mode matching implementation. The modes of this geometry are relatively trivial to define, however,

their normalization and proof of completeness is not ; the reader should consult the book of Schevchenko[20] for more information on these issues. Following this, definition of a scalar product of the orthonormal basis functions provides the mode normalization constants, proof of mutual mode orthogonality and the coupling coefficients between the surface modes of different regions (Sec. II B). Sections II C and II D implement the actual mode matching method, deriving the necessary coupled mode equations and describing a numerical approach to solve the problem using scattering matrices.

A. Normal mode fields.

Consider the geometry in Fig. 2 (a) consisting of two material half spaces, one of which is metallic, supporting surface waves that propagate in the $\pm\hat{x}$ and $\hat{y}$ directions. The geometry is taken as invariant in the $\hat{y}$ direction. The following derivations are for the complete set of orthogonal surface waves for the Transverse Magnetic (TM) polarization only; in particular, the current study is primarily concerned with the reflection, transmission and scattering of surface guided modes of the TM polarization. It is important to note here that in cases where the incident field is not invariant in the $\hat{y}$ direction (i.e. for surface waves impinging on a surface discontinuity at an angle) inter-polarization coupling occurs and including the TE radiation modes becomes necessary (See Sec. IV D). Following Schevchenko's prescription for this problem[20], the set of surface waves consists of a single bound surface mode and a continuous set of radiation modes that form a complete orthonormal set. Note that throughout this paper the term *surface waves* refers to the complete set of bound and radiation modes. In some cases, *bound surface waves* are also referred to as surface plasmons. The derivation assumes an $\exp(-\mathrm{i}\omega t)$ time variation of the field.

Solving the wave equation independently in both regions yields a continuous set of forwards and backwards propagating plane waves in the $\pm\hat{z}$ direction, as depicted in Fig. 2 (a). For TM plane waves, the unit magnetic field vector, $\hat{\mathbf{H}}$ is

$$\hat{\mathbf{H}}_{\pm} = \pm\left\{-\frac{k_y}{V}, \frac{k_x}{V}, 0\right\} \tag{1}$$

Here, the option of sign relates to the direction of wave propagation in the $\hat{z}$ direction and $V = \sqrt{k_x^2 + k_y^2}$. Note that invariance of the structure in the $\hat{x}$ and $\hat{y}$ directions in the current calculation implies conservation of both k_x and k_y, which is also known as Snell's Law.

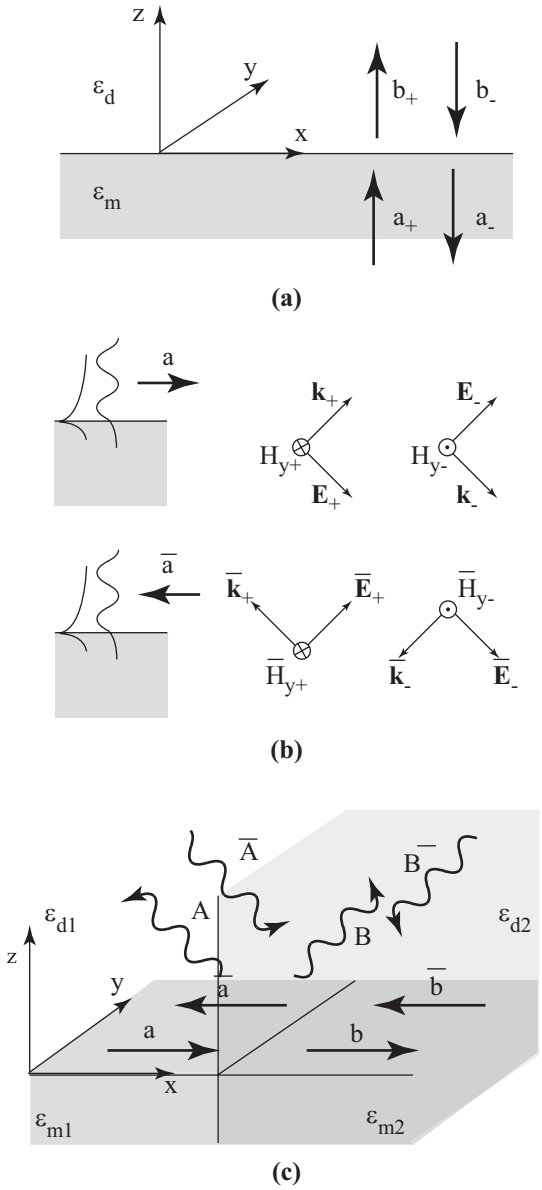

FIG. 2: (a) Schematic of the metal/dielectric open waveguide system. (b) Convention of coupled plane wave orientations for forwards and backwards propagation in both the $\hat{z}$ and $\hat{x}$ directions. (c) Schematic of the surface discontinuity problem highlighting the required modal amplitudes for both guided surface and radiating volume waves.

The total magnetic field $\mathbf{H}(\mathbf{r})$, taking into account both forward and backward propagating amplitudes is

$$\mathbf{H}(\mathbf{r}) = N\left(a_+(z)\exp\left(\mathrm{i}\mathbf{k}_+(z).\mathbf{r}\right)\hat{\mathbf{H}}_+ + a_-(z)\exp\left(\mathrm{i}\mathbf{k}_-(z).\mathbf{r}\right)\hat{\mathbf{H}}_-\right) \tag{2}$$

where N is a normalization factor to provide an orthonormal set of modes and $\mathbf{k}_\pm(z) =$

$\{k_x, k_y, \pm k_z(z)\}$. Using the Ampere-Maxwell law, $\varepsilon(z)k_0/z_0\mathbf{E}(\mathbf{r}) = \mathbf{k} \times \mathbf{H}(\mathbf{r})$, the vector electric field, $\mathbf{E}(\mathbf{r})$ is

$$\mathbf{E}(\mathbf{r}) = \frac{Nz_0}{\sqrt{\varepsilon(z)}} \left(a_+(z) \exp\left(\mathrm{i}\mathbf{k}_+(z).\mathbf{r}\right)\hat{\mathbf{E}}_+(z) + a_-(z) \exp\left(\mathrm{i}\mathbf{k}_-(z).\mathbf{r}\right)\hat{\mathbf{E}}_-(z) \right) \tag{3}$$

where the unit Electric field vector is

$$\hat{\mathbf{E}}_\pm(z) = \frac{1}{\sqrt{\varepsilon(z)}k_0} \left\{ \frac{k_x k_z(z)}{V}, \frac{k_y k_z(z)}{V}, \mp V \right\} \tag{4}$$

The complete set of modes for this geometry arises from applying the well-known boundary conditions on the transverse field components at the discontinuity. Specializing the above expressions with subscript labels m and d for metal ($z < 0$) and dielectric ($z \geq 0$) regions respectively, two transverse field matched expressions result; one as a result of the continuity of $\mathbf{H}(\mathbf{r}; z = 0)$ and the other from the continuity of $\mathbf{E}(\mathbf{r}; z = 0)$.

$$\begin{aligned} (a_+ - a_-) &= (b_+ - b_-) \\ (a_+ + a_-)\frac{k_{zm}}{\varepsilon_m} &= (b_+ + b_-)\frac{k_{zd}}{\varepsilon_d} \end{aligned} \tag{5}$$

Here, $a_\pm = a_\pm(z < 0)$ and $b_\pm = a_\pm(z \geq 0)$ as indicated in Fig. 2 (a). Since the metallic region is strongly absorbing, the analysis ignores fields propagating from $-\infty$ within the metal such that $a_+ = 0$. In this case, the full field functions are

$$\begin{aligned} \mathbf{H}(\mathbf{r}) &= -N\psi_-(z) \exp\left(\mathrm{i}k_x x + \mathrm{i}k_y y\right)\hat{\mathbf{H}}_+ \\ \mathbf{E}(\mathbf{r}) &= \frac{Nz_0}{\varepsilon(z)k_0} \left\{ \psi_+(z)\frac{k_x k_z}{V}, \psi_+(z)\frac{k_y k_z}{V}, \psi_-(z)V \right\} \exp\left(\mathrm{i}k_x x + \mathrm{i}k_y y\right) \end{aligned} \tag{6}$$

where,

$$\begin{aligned} \psi_\pm(z) &= r \exp\left(\mathrm{i}k_{zd}z\right) \pm \exp\left(-\mathrm{i}k_{zd}z\right) \quad z \geq 0 \\ \psi_\pm(z) &= \quad \pm(1 - r) \exp\left(-\mathrm{i}k_{zm}z\right) \quad z < 0 \end{aligned} \tag{7}$$

Here, the amplitudes for counter propagating plane waves in the $\hat{z}$ direction equate to: $b_- = N$, the normalization factor; and $b_+/b_- = r$, the reflectivity of the dielectric metal interface:

$$r = \frac{\varepsilon_d k_{zm} - \varepsilon_m k_{zd}}{\varepsilon_d k_{zm} + \varepsilon_m k_{zd}} \tag{8}$$

Equations (6) and (7) along with their associated dispersion relations provide the normal mode fields for this geometry describing the propagation surface waves along the $\pm\hat{x}$ and $\hat{y}$ directions. The guided mode has the following dispersion relation arising from the condition that $b_+ = 0$, irrespective of b_- and a_-: *i.e.* $r \mapsto \infty$. This provides the well known dispersion relation for surface plasmons:

$$k_{x,g} = k_0 \sqrt{\frac{\varepsilon_m \varepsilon_d}{\varepsilon_m + \varepsilon_d}} \tag{9}$$

The guided mode field functions arise from Eqns. (6) and (7) by direct substitution for $r \mapsto \infty$ and $k_x = k_{x,g}$. The dispersion relation for the radiation modes is continuous with $k_{z,r} = [0, \infty]$.

The analysis that follows examines both forwards and backwards propagating modes in both the $\pm\hat{x}$ as well as the $\pm\hat{z}$ direction giving a total of 4 plane wave solutions, which can cause some confusion with respect to the relative signs of the various field vectors. It is therefore, important to examine the relationship between the groups of counter propagating wave solutions. Figure 2 (b) highlights the conventions adopted in this paper. Here, a bar over a field vector distinguishes backwards propagating mode solutions traveling in the $-\hat{x}$ direction from individual plane waves. Direct substitution of the barred wavevector, $\bar{\mathbf{k}}_\pm(z) = \{-k_x, k_y, \pm k_z(z)\}$, provides the backwards propagating fields

$$\begin{aligned} \bar{\mathbf{H}}(\mathbf{r}) &= \{-H_x(\mathbf{r}), H_y(\mathbf{r}), 0\} \\ \bar{\mathbf{E}}(\mathbf{r}) &= \{E_x(\mathbf{r}), -E_y(\mathbf{r}), -E_z(\mathbf{r})\} \end{aligned} \tag{10}$$

In particular, the reader should note that the relative signs of counter propagating H_y and E_y field components are opposite to the counter propagating plane waves in the $\pm\hat{z}$ direction.

B. Mode orthogonality, normalization and coupling coefficients.

The complete set of surface wave solutions satisfies a mutual orthogonality condition: Collin[25] provides an introduction to the orthogonality relation shown in Eqn. (10). This integral expression also generates the coupling coefficients between the modes of two non-orthogonal eigenmode spaces. This is essentially the scalar product of the union of the two complete orthogonal eigenmode spaces. Here, Eqn. (10) shows the general expression for

the coupling coefficients between two modes of regions i and j, which naturally reduces to the mode orthogonality expression for $i = j$.

$$\begin{aligned}\langle \mathbf{E}_i(\mathbf{k}_i), \mathbf{H}_j(\mathbf{k}_j)\rangle &= \iint_{-\infty}^{\infty} \mathbf{E}_i(\mathbf{r};\mathbf{k}_i) \times \mathbf{H}_j^{\dagger}(\mathbf{r};\mathbf{k}_j).\hat{x}dydz \\ &= -\iint_{-\infty}^{\infty} E_{z,i}(\mathbf{r};\mathbf{k}_i) H_{y,j}^{\dagger}(\mathbf{r};\mathbf{k}_j)dydz \end{aligned} \tag{11}$$

Here, the dagger indicates the adjoint field, which is equivalent to reversing the propagation direction of the field[25]. The use of the adjoint field as opposed to complex conjugate, ensures hermiticity of the coupling coefficients, one o the requirements for a consistent definition of a scalar product. The requirement of linearity follows naturally from the definition. Note that both mode coupling and mode orthogonality are independent of each individual mode's harmonic variation in the $\hat{x}$ direction. Evaluating Eqn. (10) is straightforward for harmonic variation in the $\hat{y}$ direction and by direct substitution of the field components from Eqns (6) and (7), integration over y leaves an integral over z.

$$\langle \mathbf{E}_i(\mathbf{k}_i), \mathbf{H}_j(\mathbf{k}_i)\rangle = \frac{N_i(\mathbf{k}_i)N_j(\mathbf{k}_j)z_0 k_{x,j} V_i}{k_0 V_j}\delta(k_{y,i} - k_{y,j}) I(\mathbf{k}_i, \mathbf{k}_j) \tag{12}$$

where,

$$\begin{aligned} I(\mathbf{k}_i, \mathbf{k}_j) &= \int_{-\infty}^{\infty} \frac{\psi_{i,-}(z)\psi_{j,-}(z)}{\varepsilon_i(z)}dz = -\frac{(r_i + r_j)\pi}{\varepsilon_{d,i}}\delta(k_{z,i} - k_{z,j}) \\ &+ i\frac{(1 - r_i)(1 - r_j)}{\varepsilon_{m,i}(k_{z,i}^2 - k_{z,j}^2)}\left[k_{zm,i} - \frac{\varepsilon_{d,j}\varepsilon_{m,i}}{\varepsilon_{d,i}\varepsilon_{m,j}}k_{zm,j} - \frac{(k_{z,i}^2 - k_{z,j}^2)}{(k_{zm,i} + k_{zm,j})}\right] \end{aligned} \tag{13}$$

It is noteworthy that the mathematical foundation of the following mode matching analysis lies in this single integral. It provides analytical expressions for the coupling between local surface modes of various regions in a calculation. The full analytical evaluation of the integral leads to a robust and numerically stable implementation, which sets this approach aside from past reports on this subject.

In the case of radiation modes, both $k_{z,i} = [0, \infty]$ and $k_{z,j} = [0, \infty]$, so that the radiation to radiation mode coupling coefficients retain all parts of Eqn. (13) . Locally, radiation modes are orthogonal by inspection, such that Eqn. (13) evaluates to,

$$\langle \mathbf{E}_i(\mathbf{k}_i), \mathbf{H}_i((\mathbf{k}_i'))\rangle = -\pi\frac{N_i(\mathbf{k}_i)N_i(\mathbf{k}_i')(r_i + r_i')z_0 k_{x,i}'}{\varepsilon_{d_i} k_0}\delta(k_{y,i} - k_{y,i}')\delta(k_{z,i} - k_{z,i}') \tag{14}$$

Therefore, the radiation modes are orthonormal with the normalization constant,

$$N_i(\mathbf{k}_i) = \mathrm{i}\sqrt{\frac{\varepsilon_d k_0}{2\pi r_i z_0 k_{x,i}}} \tag{15}$$

Locally, radiation modes and guided modes are orthogonal by inspection of Eqn. (13). For the guided modes $r = r' \mapsto \infty$. Under these conditions, there is only one guided mode; in this case, distinguishing the guided mode's fields by small letters reduces clutter and the need to specify continuous wavevectors: therefore, $\{\mathbf{E}_i(\mathbf{k}_i; k_x = k_{xg,i}), \mathbf{H}_j(\mathbf{k}_i; k_x = k_{xg,i})\} = \{\mathbf{e}_i, \mathbf{h}_j\}$ and equivalently for all parameters that are function of $\mathbf{k}_i$. Locally, their is no orthogonality condition as there is one guided mode such that,

$$\langle \mathbf{e}_i, \mathbf{h}_i \rangle = \mathrm{i}\frac{N_{g,i}^2 z_0 k_{x,i}}{k_0}\left[\frac{\varepsilon_d k_{zd,i} - \varepsilon_{m,i} k_{zm,i}}{\varepsilon_{d,i}\varepsilon_{m,i} k_{zd,i} k_{zm,i}}\right] \tag{16}$$

The guided modes are therefore orthonormal with the normalization constant,

$$N_{g,i} = -\sqrt{\frac{k_0}{z_0 k_{x,i}}}\left[\frac{2\mathrm{i}\varepsilon_{d,i}\varepsilon_{m,i}^2 k_{zd,i}}{\varepsilon_{d,i}^2 - \varepsilon_{m,i}^2}\right]^{1/2} \tag{17}$$

Notice, in the guided mode case, that the factor $(1 - r_i)(1 - r_j) \mapsto \infty$ from the coupling coefficient expression of Eqn. (13) is missing from the definition of $N_{g,i}$. In mode matching studies, it is the shape of the mode that is relevant and $N_{g,i}$ eliminates the singularity as $r_i \mapsto \infty$, although Eqn. (17) does not explicitly specify it. A consequence of removing the singularity requires an additional minus sign in the definition of $N_{g,i}$: this is because the terms $(1 - r_i)$ and $(1 - r_j)$ both tend to $-\infty$ separately.

With the complete set of orthonormal modes for the local geometry in Fig. 2 (a), it is possible to consistently relate the field on either side of a surface discontinuity at $x = 0$ through the evaluation of coupling coefficients. These arise by direct substitution of the suitably normalized field components into Eqn. (12). Correct mode matching requires the following generally non-zero coefficients: radiation to radiation - $\langle \mathbf{E}_i(\mathbf{k}_i), \mathbf{H}_i(\mathbf{k}_j) \rangle$; guided to guided - $\langle \mathbf{e}_i, \mathbf{h}_j \rangle$; guided to radiation $\langle \mathbf{e}_i, \mathbf{H}_j(\mathbf{k}_j) \rangle$; and radiation to guided $\langle \mathbf{E}_i(\mathbf{k}_i), \mathbf{h}_j \rangle$. In addition, when using normalized fields, as is the case here, the rest of the coupling coefficients simplify greatly for mode coupling within the same region: $\langle \mathbf{E}_i(\mathbf{k}_i), \mathbf{H}_i(\mathbf{k}_i') \rangle = \delta(k_{y,i} - k_{y,i}')\delta(k_{z,i} - k_{z,i}')$, $\langle \mathbf{e}_i, \mathbf{h}_i \rangle = 1$ and $\langle \mathbf{e}_i, \mathbf{H}_i(\mathbf{k}_i) \rangle = \langle \mathbf{E}_i(\mathbf{k}_i), \mathbf{h}_i \rangle = 0$.

C. Mode matching equations

The mode matching equations arise from the matching of field components across the discontinuity at $x = 0$, shown in Fig. 2 (c) followed by use of the coupling coefficients

integral of Eqn. (12). Using the expressions from Sec. II B, the continuity of the transverse components of the electric, E_z, and magnetic, H_y, fields is

$$\begin{aligned}
&(a-\bar{a})\, e_{z,i}(\mathbf{r}) + \int_0^\infty \left(A(\mathbf{k}) - \bar{A}(\mathbf{k})\right) E_{z,i}(\mathbf{r};\mathbf{k}) dk_z \\
&= \left(b-\bar{b}\right) e_{z,j}(\mathbf{r}) + \int_0^\infty \left(B(\mathbf{k}) - \bar{B}(\mathbf{k})\right) E_{z,j}(\mathbf{r};\mathbf{k}) dk_z \qquad (18) \\
&(a+\bar{a})\, h_{y,i}(\mathbf{r}) + \int_0^\infty \left(A(\mathbf{k}) + \bar{A}(\mathbf{k})\right) H_{y,i}(\mathbf{r};\mathbf{k}) dk_z \\
&= \left(b+\bar{b}\right) h_{y,j}(\mathbf{r}) + \int_0^\infty \left(B(\mathbf{k}) + \bar{B}(\mathbf{k})\right) H_{y,j}(\mathbf{r};\mathbf{k}) dk_z \qquad (19)
\end{aligned}$$

This is just a summation over guided and radiation modes with integration over the continuous radiation modes. Choosing to integrate the radiation modes with respect to k_z proves to be useful later on when calculating the specular coupling of radiation modes. Post vector multiplication of Eqn. (18) by $\times h_{y,i}(\mathbf{r})$ and $\times H_{y,i}(\mathbf{r})$ followed by vector integration over the z, y plane provides 2 mode matching equations. Similarly, pre-vector multiplication of Eqn. (19) by $e_{z,j}(\mathbf{r})\times$ and $E_{z,j}(\mathbf{r})\times$ followed by vector integration over the z, y plane provides the other 2.

$$\begin{aligned}
a-\bar{a} &= \left(b-\bar{b}\right)\langle \mathbf{e}_j, \mathbf{h}_i\rangle + \int_0^\infty \left(B(\mathbf{k}') - \bar{B}(\mathbf{k}')\right)\langle \mathbf{E}_j(\mathbf{k}'), \mathbf{h}_i\rangle\, dk_z' \\
A(\mathbf{k}) - \bar{A}(\mathbf{k}) &= \left(b-\bar{b}\right)\langle \mathbf{e}_j, \mathbf{H}_i(\mathbf{k})\rangle + \int_0^\infty \left(B(\mathbf{k}') - \bar{B}(\mathbf{k}')\right)\langle \mathbf{E}_j(\mathbf{k}), \mathbf{H}_i(\mathbf{k}')\rangle\, dk_z' \\
b+\bar{b} &= (a+\bar{a})\langle \mathbf{e}_j, \mathbf{h}_i\rangle + \int_0^\infty \left(A(\mathbf{k}') + \bar{A}(\mathbf{k}')\right)\langle \mathbf{e}_j, \mathbf{H}_i(\mathbf{k}')\rangle\, dk_z' \\
B(\mathbf{k}) + \bar{B}(\mathbf{k}) &= (a+\bar{a})\langle \mathbf{E}_j(\mathbf{k}), \mathbf{h}_i\rangle + \int_0^\infty \left(A(\mathbf{k}') + \bar{A}(\mathbf{k}')\right)\langle \mathbf{E}_j(\mathbf{k}), \mathbf{H}_i(\mathbf{k})\rangle\, dk_z' \qquad (20)
\end{aligned}$$

In line with the mode matching method, truncated summations approximate the integrals over the continuous set of radiation modes to leave a matrix formulation. The current implementation is slightly different from traditional approaches in that the integrals are written as a Gaussian Quadrature summation. In this case, the quadrature weighting factors require special treatment: the solution to the mode matching problem yields a set of radiation modes that, when summed, gives the total scattered power. So, for example, in the case of backscattering, the truncated summation approximation for the integral is,

$$\int |\bar{A}(x;\mathbf{k})|^2 . dk_z \mapsto \sum_m |\bar{A}_m(x)|^2 w_m \qquad (21)$$

It follows that $A(x;\mathbf{k}) \mapsto A_m/\sqrt{w_m}$ provides the correct transformation between the continuous and discrete representations, where w_m are the weighting coefficients of the Gaussian

Quadrature scheme. Approximating the integrals in Eqn. (20) with Gaussian Quadrature summations and substituting for the normalized radiation mode amplitudes, the following mode matching equations result

$$\begin{aligned}
a - \bar{a} &= \left(b - \bar{b}\right) \langle \mathbf{e}_j, \mathbf{h}_i \rangle + \sum_m \left(B_m - \bar{B}_m\right) \langle \mathbf{E}_j(\mathbf{k}_m), \mathbf{h}_i \rangle \sqrt{w_m} \\
A_n - \bar{A}_n &= \left(b - \bar{b}\right) \langle \mathbf{e}_j, \mathbf{H}_i(\mathbf{k}_n) \rangle \sqrt{w_m} + \left(B_n - \bar{B}_n\right) \langle \mathbf{E}_j(\mathbf{k}_n), \mathbf{H}_i(\mathbf{k}_n) \rangle \\
&\quad + \sum_m \left(B_m - \bar{B}_m\right) \langle \mathbf{E}_j(\mathbf{k}_m), \mathbf{H}_i(\mathbf{k}_n) \rangle \, w_m \\
b + \bar{b} &= (a + \bar{a}) \langle \mathbf{e}_j, \mathbf{h}_i \rangle + \sum_m \left(A_m + \bar{A}_m\right) \langle \mathbf{e}_j, \mathbf{H}_i(\mathbf{k}_m) \rangle \sqrt{w_m} \\
B_n + \bar{B}_n &= (a + \bar{a}) \langle \mathbf{E}_j(\mathbf{k}_n), \mathbf{h}_i \rangle \sqrt{w_m} + \left(A_n + \bar{A}_n\right) \langle \mathbf{E}_j(\mathbf{k}_n), \mathbf{H}_i(\mathbf{k}_n) \rangle \\
&\quad + \sum_m \left(A_m + \bar{A}_m\right) \langle \mathbf{E}_j(\mathbf{k}_n), \mathbf{H}_i(\mathbf{k}_m) \rangle \, w_m
\end{aligned} \tag{22}$$

Arranging these expressions into two coupled matrix equations:

$$\begin{aligned}
A_n + \bar{A}_n &= \sum_m \left(B_m + \bar{B}_m\right) C_{mn} \\
B_n - \bar{B}_n &= \sum_m \left(A_m - \bar{A}_m\right) C^T_{mn}
\end{aligned} \tag{23}$$

where,

$$\begin{aligned}
C_{00} &= \langle \mathbf{e}_j, \mathbf{h}_i \rangle \;\; ; \;\; m = n = 0 \\
C_{0n} &= \langle \mathbf{e}_j, \mathbf{H}_i(\mathbf{k}_n) \rangle \sqrt{w_m} \\
C_{m0} &= \langle \mathbf{E}_j(\mathbf{k}_m), \mathbf{h}_i \rangle \sqrt{w_m} \\
C_{mn} &= \langle \mathbf{E}_j(\mathbf{k}_m), \mathbf{H}_i(\mathbf{k}_n) \rangle \;\; ; \;\; m = n \\
C_{mn} &= \langle \mathbf{E}_j(\mathbf{k}_m), \mathbf{H}_i(\mathbf{k}_n) \rangle \, w_m \;\; ; \;\; m \neq n
\end{aligned} \tag{24}$$

It is important to note that the elements C_{mn} are not trivial to evaluate analytically due to the singularity at $k_{zd,i} = k_{zd,j}$. In this case, only the non-singular components of Eqn. (12) contribute to C_{mn}; the remaining term in $(k_{zd,i} - k_{zd,j})^{-1}$ contributes very little as it's principal value in the vicinity of the singularity is negligible. The weight functions in Eqn. (24) indicate the strength of coupling between the guided and radiation modes. Guided to radiation mode coupling is clearly important with a scaling of $\sqrt{w_m}$. Whereas, scattering between un-matched radiation modes is not as important with a scaling w_m. Again, examining the parts of the C_{mn} elements from Eqn. (12) shows that the contribution from the delta function is the most significant term since the other parts scale with w_m.

Manipulating the matrix equations further gives the interface scattering matrix relating the modal outputs to the inputs such that,

$$\begin{pmatrix} \mathbf{B} \\ \bar{\mathbf{A}} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} \begin{pmatrix} \bar{\mathbf{B}} \\ \mathbf{A} \end{pmatrix} \tag{25}$$

where,

$$\begin{aligned} \mathbf{S}_{11} &= \left[\mathbf{1} + \mathbf{C}\mathbf{C}^T\right]^{-1} \left(\mathbf{1} - \mathbf{C}\mathbf{C}^T\right) \\ \mathbf{S}_{12} &= \left[\mathbf{1} + \mathbf{C}\mathbf{C}^T\right]^{-1} \mathbf{C} \\ \mathbf{S}_{21} &= \left[\mathbf{1} + \mathbf{C}^T\mathbf{C}\right]^{-1} \mathbf{C}^T \\ \mathbf{S}_{22} &= \left[\mathbf{1} + \mathbf{C}\mathbf{C}^T\right]^{-1} \left(\mathbf{1} - \mathbf{C}\mathbf{C}^T\right) \end{aligned} \tag{26}$$

D. Multiple Interface Calculations

The scattering matrix method describes general multiple surface discontinuities by defining a propagation scattering matrix for the discretized system of modes from Sec. II C. The scattering matrix relating amplitudes within the same region propagating between the positions x_1 and x_2 has the form:

$$\begin{pmatrix} \mathbf{A}(x_2) \\ \bar{\mathbf{A}}(x_1) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{P} \\ \mathbf{P} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \bar{\mathbf{A}}(x_2) \\ \mathbf{A}(x_1) \end{pmatrix} \tag{27}$$

where,

$$\begin{aligned} \mathbf{P}_{mn} &= \exp\left\{\mathrm{i}k_{xg}(x_2 - x_1)\right\} \quad m = n = 0 \\ &= \exp\left\{\mathrm{i}k_{xr,m}(x_2 - x_1)\right\}\delta_{mn} \end{aligned} \tag{28}$$

Combining both interface (Eqn. (26)) and propagation (Eqn. (27)) scattering matrices provides the description of any system of co-planar discontinuities of the form under investigation. Note that combining scattering matrices requires the usual concatenation method.

III. NUMERICAL ANALYSIS

This section assesses the accuracy and convergence of the modeling method with a series of numerical tests. Firstly, checking the matching of field functions on either side of an

open waveguide discontinuity ensures both correct operation of the model as well as self-consistency. Secondly, the method is compared with the commercial finite element modeling software from Comsol (FEMLab).

Firstly, consider a comparison of the matched fields at the abrupt discontinuity of two open waveguides with $\varepsilon_{m,1} = \varepsilon_{m,2} = -18.3 - 0.5i$ (ε_{Ag} at $\lambda = 632.8$ nm[27]), $\varepsilon_{d,1} = 2.25$ and $\varepsilon_{d,2} = 1$. To generate an accurate matching of field functions, the model uses 200 radiation modes in addition to the single surface plasmon mode. Studies show that solution convergence requires a large number of evanescent modes; here, the radiation mode truncation is set at $k_{z,max} = 10\sqrt{\varepsilon_{d,max}}k_0$, where $\varepsilon_{d,max}$ is the highest dielectric permittivity in a structure. This prescription works well, although small improvements in convergence occur with $k_{z,max} = 100\sqrt{\varepsilon_{d,max}}k_0$. Fig. 3 (a) and 3 (b) show the real and imaginary parts of the H_y and E_z fields respectively on either side of the discontinuity; clearly, the correspondence is excellent. A more detailed examination of the fields very near the metal dielectric interface highlights the largest deviations from ideal field matching. Using more radiation modes, a larger proportion of which are evanescent, eliminates these errors indicating asymptotic convergence to the true solution. The mean standard deviation between the fields on either side of the discontinuity for the case considered above is slightly larger than 1 %. With 500 radiation modes truncated at $k_{z,max} = 50\sqrt{\varepsilon_{d,max}}k_0$, the deviation is 0.5 %. A high level of convergence clearly requires significant computation expenditure as noted originally for this problem by Stegeman and co-workers[22]. Note that field matching requires more accuracy: satisfactory convergence of the modal amplitudes occurs for about 50 modes.

Upon comparison with Comsol's finite element software for an identical problem, the correspondence is also remarkable. Figures 4 (a) to 4 (c) compare the absolute field components $|H_y|$, $|E_z|$ and $|D_x|$ respectively at the same open waveguide discontinuity as in Fig. 3 for the finite element (solid lines) and semi-analytical (broken lines) methods. The FEM software employed scattering boundary conditions to minimize interference between the solution and boundary reflections within the calculation domain: this was a rectangle $10\,\mu$m (x-direction) by $2\,\mu$m (y-direction), split into equal quarters of permittivity $\varepsilon_{d,1}$, $\varepsilon_{m,1}$, $\varepsilon_{m,2}$ and $\varepsilon_{d,2}$ anti-clockwise from the top left quarter. Convergence of the FEM software for this problem required a pre-generated adaptive mesh with nearly $500,000$ elements and a calculation time of approximately 2 mins. on a 2.4 GHz Machine. The semi-analytical result took only 5 s using MatLab on the same machine. Note that a direct comparison of the

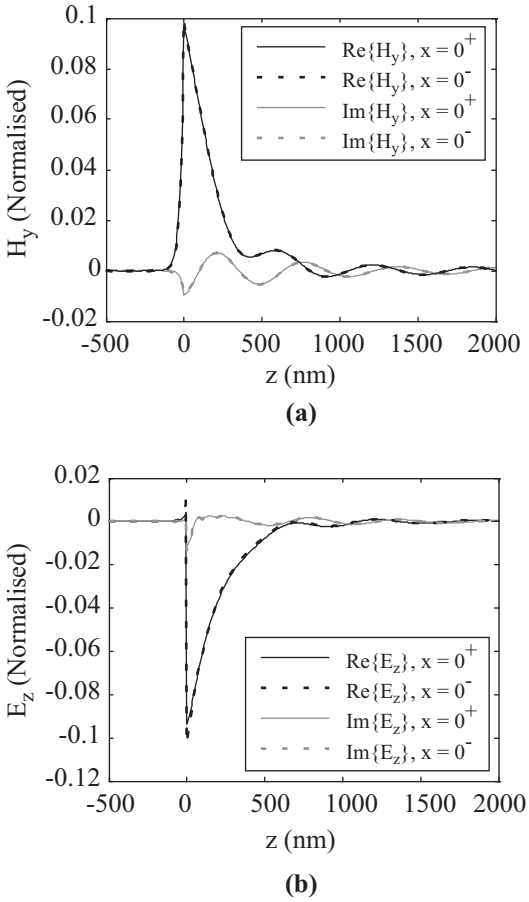

FIG. 3: Comparison of the matched (a) H_y and (b) E_z fields on either side ($x = 0^+$, 0^-) of an abrupt open waveguide discontinuity. Here, $\varepsilon_{m,1} = \varepsilon_{m,2} = -18.3 - 0.5i$ (ε_{Ag} at $\lambda = 632.8$ nm), $\varepsilon_{d,1} = 2.25$ and $\varepsilon_{d,2} = 1$.

two methods requires normalization of both results to the same peak height. The accuracy of either method is clearly not in question from the results in Figs. 3 and 4. Although the time saving in using the semi-analytical approach is a clear advantage, the principal benefit of this approach is that it provides the modal amplitudes directly. In fact, in generating the field plots in Figs. 3 and 4, the analytical approach must perform additional computations that constitute almost 50 % of the overall calculation time. In stark contrast, de-convolving the modal scattering amplitudes from the fields of the finite element approach is not a trivial task.

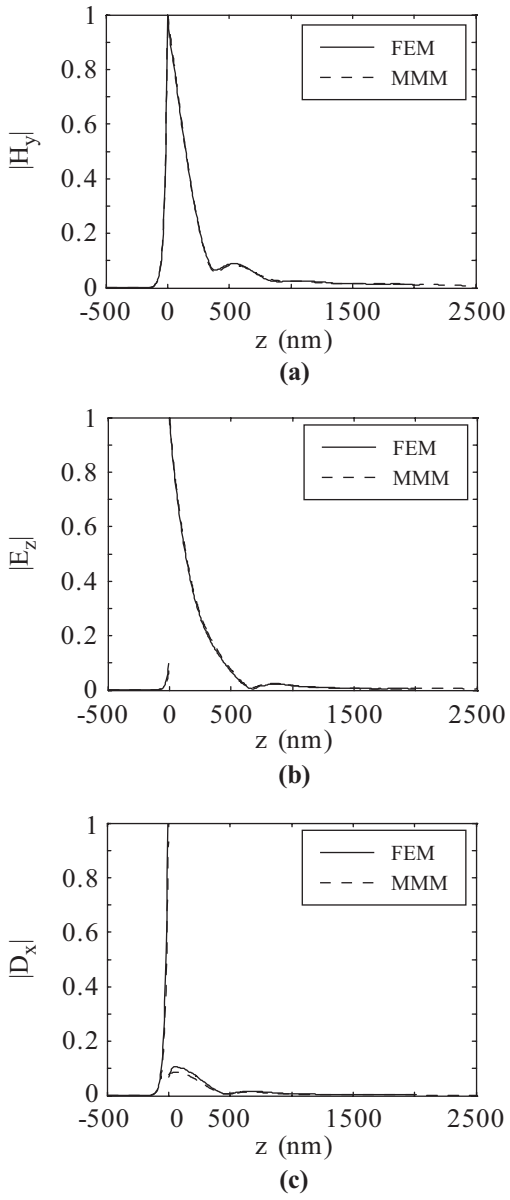

FIG. 4: Comparison of the matched (a) $|H_y|$ and (b) $|D_x|$ and (c) $|E_z|$ fields at an abrupt open waveguide discontinuity with FEMLab simulations. Here, $\varepsilon_{m,1} = \varepsilon_{m,2} = -18.3 - 0.5i$ (ε_{Ag} at $\lambda = 632.8$ nm), $\varepsilon_{d,1} = 2.25$ and $\varepsilon_{d,2} = 1$

IV. RESULTS

The results section is split into four parts. The first two parts show the results of transmission, reflection and scattering calculations for a surface plasmon wave incident on an interface between two dielectric / metal half spaces; the first part considers dielectric discontinuities and the second part considers metallic discontinuities. The third part considers

the compilation of multiple surface discontinuities into a surface plasmon Distributed Bragg Reflector (DBR). The final part introduces full the 3D calculations for surface plasmon waves incident at an angle to a surface discontinuity. In all cases, the results show the scattering effect of a single incident surface plasmon wave. Figure 2 (c) shows the general case of surface waves that enter and leave a region of surface discontinuities. In the following calculations, a surface wave impinges from the left such that $a = 1$, $\bar{b} = 0$, $\bar{a}$ is the surface plasmon reflectivity and $\bar{b}$ is the surface plasmon transmission. Radiation modes propagating toward a discontinuity are set to zero, $A = \bar{B} = 0$, such that $\bar{A}$ and B quantify backward and forward scattering respectively.

In all the calculations, a single surface plasmon mode with 200 radiation modes, truncated at $k_{z,max} = 10\sqrt{\varepsilon_{d,max}}k_0$, where $\varepsilon_{d,max}$ is the maximum permittivity under consideration, constitute the numerical eigenvector space. The reader is referred to the numerical analysis in Sec. III for more detail. When not scanning spectrally, results of the following study are for the Helium-Neon red laser wavelength at 632.8 nm.

A. Dielectric discontinuity.

As previously noted in Sec. III, the mode matching method directly generates modal scattering amplitudes. Figure 5 shows the reflection, transmission and scattering to radiation modes for a surface plasmon wave normally incident on a dielectric discontinuity, yet guided along the same continuous Silver surface. Note that the proportion of forwards and backwards scattered radiation corresponds to an integration over the propagating radiation modes in the respective medium. The substantial evanescent wave angular spectrum, which is important in the case of multiple discontinuities, does not contribute to the scattering loss here. The interface calculations satisfy energy conservation by virtue of the fact that $S \approx 1 - R - T$ to an acceptable tolerance; note that for the high to low permittivity case of Fig. 5 (b) would converge better with more radiation modes and a higher evanescent wavevector truncation.

A particularly interesting feature of these calculation is the close correspondence of the surface plasmon reflectivity with the Fresnel value that accounts for the effective phase index

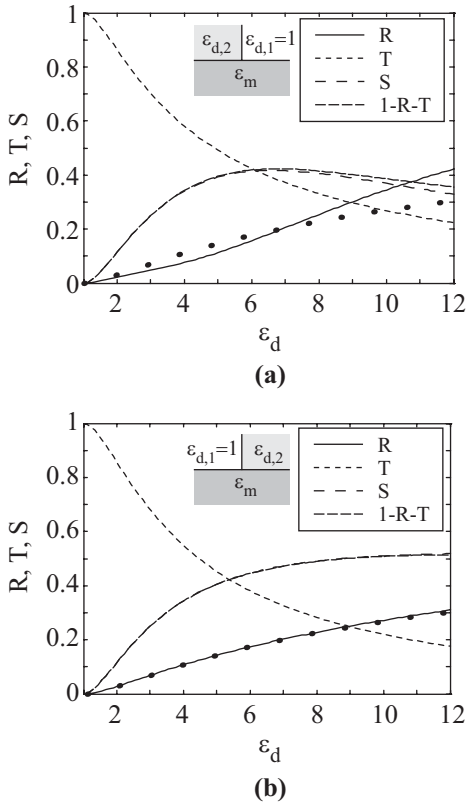

FIG. 5: Reflection, transmission and scattering of surface waves (propagating left to right) as a function of dielectric permittivity contrast $\varepsilon_{d,1}$ and $\varepsilon_{d,2}$ for (a) high to low and (b) low to high at a wavelength of 632.8 nm. Dots show the equivalent Fresnel reflection coefficient calculated from Eqn. 29.

of the surface wave, n_i. The dots indicate the Fresnel reflectivity, R_F, given by,

$$R_F = \left| \frac{(n_i - n_j)}{(n_i + n_j)} \right|^2 \tag{29}$$

where $n_i = \sqrt{\varepsilon_{d,i}\varepsilon_{m,i}/(\varepsilon_{d,i} + \varepsilon_{m,i})}$. The correspondence is exact for low to high surface plasmon wavevectors, with a small deviation for the high to low permittivity case. Related reports have identified similar non-reciprocal behavior[24]. Analysis of the radiation mode distributions, shown in Fig. 6 for $\varepsilon_{d,i} = 2.25$ and $\varepsilon_{d,j} = 1$ interfaces with Silver, reveals further non-reciprocal behavior. In both cases most of the scattered power is in the forward direction, however, only in the case of high to low permittivity is there any significant back scattering. This provides an indication of the source of a deviation from the Fresnel coefficient. These results are in basic qualitative agreement with those of Voronko[23], however, quantitatively, there are significant discrepancies. It is noteworthy that, as a rule of thumb,

the reflection is equivalent to the Fresnel value for bulk waves and the transmission and scattering share from the remaining electromagnetic power with proportions that depend on the size of the discontinuity.

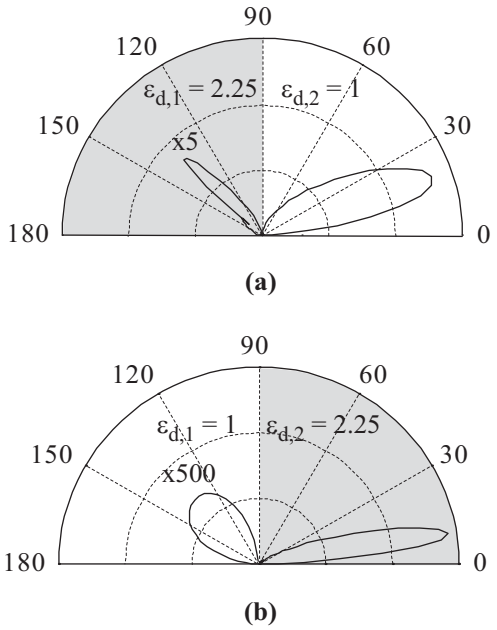

FIG. 6: Scattering of propagating radiation waves at the dielectric / metal surface discontinuity for (a) $\varepsilon_{d,1} = 2.25$ to $\varepsilon_{d,2} = 1$ on Silver, and (b) $\varepsilon_{d,1} = 1$ to $\varepsilon_{d,2} = 2.25$ on Silver at a wavelength of 632.8 nm.

B. Metallic discontinuity

Consider now the scattering of surface plasmon waves normally incident on a metallic surface discontinuity, where the dielectric half space is continuous. This study briefly outlines the effects of a perfect abrupt transition between two metallic half spaces, although in practice such a structure may be complicated to realize. Figure 7 shows results of calculating the transition between regions of Silver/Air and Aluminium/Air. The calculations consider real permittivity data for both Aluminium[26] and Silver[27].

Surface plasmon reflection is minimal (< 1 %) and is only significant near the surface plasma edge of Silver, where the mode shape and effective index of the modes on either side of the discontinuity most different. The small differences in the shape of surface plasmon modes in each region causes a significant amount of scattering, settling to about 10 % far

from the surface plasmon edge. As with the case of the dielectric discontinuity, forward scattering is dominant with significant non-reciprocal backwards scattering for the Silver to Aluminium and Aluminium to Silver cases. Again, these results are in qualitative agreement with those of Voronko[23]. Following a similar approach to the dielectric case in Sec. IV A, the

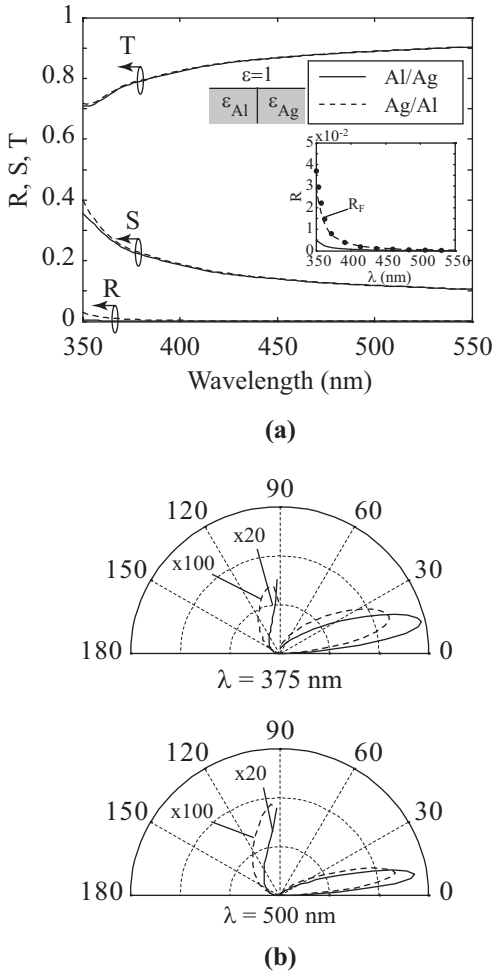

FIG. 7: Surface plasmon scattering at a metallic discontinuity. (a) reflection, transmission and scattering coefficients. (b) identifies the direction of radiated scattering. Solid lines represent transition from Aluminium to Silver and the broken lines represent the reverse. The inset of (a) shows the reflection coefficients compared with effective Fresnel reflection calculated from Eqn. (29).

inset of Fig. 7 (a) compares the reflection coefficients with the effective Fresnel coefficient calculated using Eqn. (29): the effective index for a surface plasmon on Silver in the current case is $n_{Ag}(\lambda) = \sqrt{\varepsilon_{Ag}(\lambda)/(\varepsilon_{Ag}(\lambda)+1)}$ (since $\varepsilon_d = 1$) and a similar expression holds for $n_{Al}(\lambda)$. Again, the effective reflection coefficient describes the case of Silver to Aluminium

well. However, the reflectivity of this system is even more asymmetric than in the dielectric case: for Aluminium to Silver, the results do not closely match at all. Again, the strength of backward scattering in this case is 5 times higher and indicates that the discrepancy arises due to more significant modal mismatch.

The various combinations of Gold, Silver and Aluminium produce similar results, with moderately significant reflection only manifested near to a surface plasmon edge. It seems, therefore, that this system is unappealing for creating resonant systems: given the intrinsic damping of surface waves, strong reflections at discontinuities would be more effective in a distributed style reflector. As the wavelength increases in these systems, the surface plasmon wave dispersion approaches the light line and therefore takes on more of the characteristics of a bulk wave. Since there is no dielectric discontinuity, scattering and reflection will settle to an asymptotic value determined by the relative conductivity of the two metals.

C. Surface Plasmon Distributed Bragg Reflector.

The final example of the current mode matching approach examines the potential to design effective surface wave resonators. One of the key components of nano and micro scale cavities are highly reflective mirrors. Distributed Bragg Reflectors (DBRs) are widely used in microcavity physics to create effective mirrors for laser devices. The possibility to use such structures on the nano scale is extremely appealing in order to enhance the low Q factors of nano-particulate systems. Figure 8 plots the results of a mode matching analysis of DBRs composed of alternating dielectric layers of PMMA ($\varepsilon_d = 2.25$) and Air next to a Silver metallic interface. The optical thickness of each layer is a quarter of the surface plasmon wavelength at $\lambda_{DBR} = 550$ nm. Under these conditions, the DBR is mainly effective for surface plasmons and a few radiation modes near the light line. Launching a surface plasmon from one side of the structure allows calculation of the effectiveness of a DBR to reflect and transmit surface plasmons as well as scatter radiation waves. The first important feature of the results in Fig. 8 (a) is that only moderate surface plasmon reflection, $R \approx 80$ %, are achievable in this type of structure. For a single wavelength cavity formed between two such reflectors, the Q-factor could approach $Q = 2\lambda_{SP}/\lambda\sqrt{R}/(1-R) \sim 10$. Figure 8 (b) highlights that the limitations are solely from scattering to radiation modes. As the number of DBR periods increases, the scattering only moderately increases, beyond

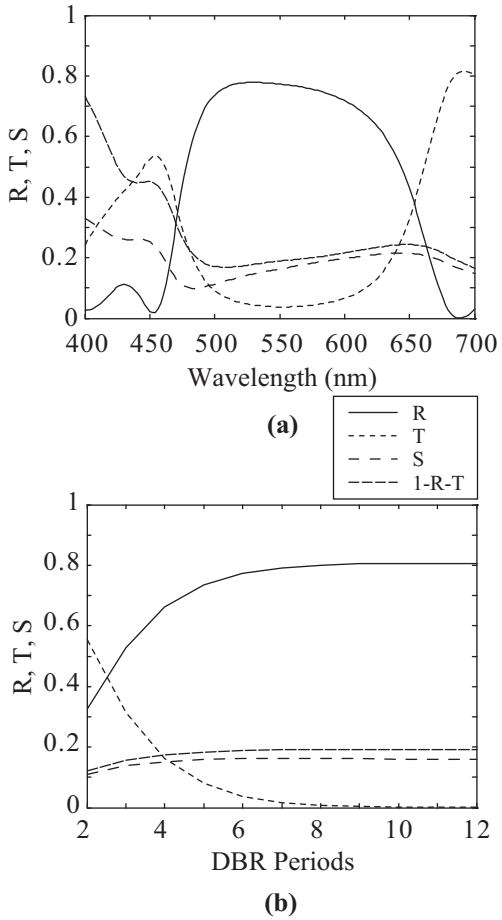

FIG. 8: Surface plasmon reflection, transmission and radiation for a PMMA / Air DBRs. (a) Spectral response of modal scattering parameters for a 6 period DBR. (b) Peak DBR reflectivity as function of the number of periods.

the 6 period result. The remaining power is shared between surface plasmon transmission and reflection through the DBR. Eventually, the reflectivity saturates and the transmission tends to zero. Although omitted from this paper, when optimizing this system to maximize reflectivity with lower permittivity contrasts brings moderate improvements of $R = 85$ %; this requires balancing intrinsic propagation losses for a given number of periods with the scattering loss of each interface. However, one can see from Fig. 8 (a) that the scattered power is larger that that due to intrinsic losses (by comparing $1 - R - T$ with the scattered power): in this case, propagation losses comprise about 10 % of the scattering losses. This is encouraging since there are other surface plasmon systems that suppress radiation scattering effectively such as in coupled plasmons between the dielectric gap of two co-planar metal

interfaces[28,29].

D. Extension to 3D calculations

Throughout this paper, derivations have involved general TM wave propagation in all 3 spatial dimensions although, up until now, the analysis has focused solely on normally incident SP waves. In principle, taking k_y as non-zero in the mode matching equations is sufficient to describe the behavior of SP waves impinging on surface discontinuities and at an angle from the normal within the $\hat{\mathbf{x}}$ - $\hat{\mathbf{y}}$ plane, however, care must taken here, as in the general 3D geometry of Fig. 2 (c) the TE (s) and TM (p) polarizations may scatter off each other. Consider the following unit field vectors for the TE radiation modes,

$$\hat{\mathbf{E}}_{\pm}^{(s)} = \left\{ -\frac{k_y}{V}, \frac{k_x}{V}, 0 \right\} \tag{30}$$

$$\hat{\mathbf{H}}_{\pm}^{(s)}(z) = \pm \frac{1}{\sqrt{\varepsilon(z)} k_0} \left\{ -\frac{k_x k_z(z)}{V}, -\frac{k_y k_z(z)}{V}, \pm V \right\} \tag{31}$$

It is immediately apparent that the field continuity expressions of Eqn. (18) require alteration for the general 3D calculation to incorporate the coupling of TE waves through the non-zero $H_{y,i}^{s}(\mathbf{r};\mathbf{k})$ component.

$$\begin{aligned}
(a_p - \bar{a}_p)\, e_{z,i}^{(p)}(\mathbf{r}) \;+\; & \int_0^\infty \left(A_p(\mathbf{k}) - \bar{A}_p(\mathbf{k}) \right) E_{z,i}^{(p)}(\mathbf{r};\mathbf{k}) dk_z \\
= \; & \left(b_p - \bar{b}_p \right) e_{z,j}^{(p)}(\mathbf{r}) + \int_0^\infty \left(B_p(\mathbf{k}) - \bar{B}_p(\mathbf{k}) \right) E_{z,j}^{(p)}(\mathbf{r};\mathbf{k}) dk_z \qquad (32) \\
(a_p + \bar{a}_p)\, h_{y,i}^{(p)}(\mathbf{r}) \;+\; & \int_0^\infty \left(A_p(\mathbf{k}) + \bar{A}_p(\mathbf{k}) \right) H_{y,i}^{(p)}(\mathbf{r};\mathbf{k}) dk_z \\
+ \; & \int_0^\infty \left(A_s(\mathbf{k}) + \bar{A}_s(\mathbf{k}) \right) H_{y,i}^{(s)}(\mathbf{r};\mathbf{k}) dk_z \\
= \; & \left(b + \bar{b} \right) h_{y,j}^{(p)}(\mathbf{r}) + \int_0^\infty \left(B_p(\mathbf{k}) + \bar{B}_p(\mathbf{k}) \right) H_{y,j}^{(p)}(\mathbf{r};\mathbf{k}) dk_z \\
+ \; & \int_0^\infty \left(B_s(\mathbf{k}) + \bar{B}_s(\mathbf{k}) \right) H_{y,j}^{(s)}(\mathbf{r};\mathbf{k}) dk_z \qquad (33)
\end{aligned}$$

Since $E_{z,i}^{s}(\mathbf{r};\mathbf{k}) = 0$, the first continuity expression remains unaltered. This moderately increases the complexity, as the reader will notice that the coupling coefficients $\left\langle \mathbf{E}_i^{(s)}(\mathbf{k}_i), \mathbf{H}_j^{(p)}(\mathbf{k}_j') \right\rangle = \left\langle \mathbf{E}_i^{(s)}(\mathbf{k}_i), \mathbf{h}_j^{(p)} \right\rangle = 0$ in addition to the usual local mode orthogonality which provide significant simplifications. By field matching the E_y, H_y, E_z and H_z components, a set of 6 mode matching equations arise, which uniquely express 2 unknown

surface plasmon amplitudes, b, $\bar{b}$, 2 unknown sets of continuous TM radiation modes, $B^{(p)}(\mathbf{k})$, $\bar{B}^{(p)}(\mathbf{k})$, and another 2 unknown sets of continuous TE radiation modes, $B^{(s)}(\mathbf{k})$, $\bar{B}^{(s)}(\mathbf{k})$ assuming *a priori* knowledge of the input amplitudes, a, $\bar{a}$, $A^{(p)}(\mathbf{k})$, $\bar{A}^{(p)}(\mathbf{k})$, $A^{(s)}(\mathbf{k})$ and $\bar{A}^{(s)}(\mathbf{k})$.

$$
\begin{aligned}
a_p - \bar{a}_p &= \left(b_p - \bar{b}_p\right)\left\langle \mathbf{e}_j^{(p)}, \mathbf{h}_i^{(p)} \right\rangle + \int_0^\infty \left(B_p(\mathbf{k}') - \bar{B}_p(\mathbf{k}')\right)\left\langle \mathbf{E}_j^{(p)}(\mathbf{k}'), \mathbf{h}_i^{(p)} \right\rangle dk_z' \\
A_p(\mathbf{k}) - \bar{A}_p(\mathbf{k}) &= \left(b_p - \bar{b}_p\right)\left\langle \mathbf{e}_j^{(p)}, \mathbf{H}_i^{(p)}(\mathbf{k}) \right\rangle + \int_0^\infty \left(B_p(\mathbf{k}') - \bar{B}_p(\mathbf{k}')\right)\left\langle \mathbf{E}_j^{(p)}(\mathbf{k}'), \mathbf{H}_i^{(p)}(\mathbf{k}) \right\rangle dk_z' \\
A_s(\mathbf{k}) - \bar{A}_s(\mathbf{k}) &= \int_0^\infty \left(B_s(\mathbf{k}') - \bar{B}_s(\mathbf{k}')\right)\left\langle \mathbf{E}_i^{(s)}(\mathbf{k}), \mathbf{H}_j^{(s)}(\mathbf{k}') \right\rangle dk_z' \\
b_p + \bar{b}_p &= (a_p + \bar{a}_p)\left\langle \mathbf{e}_j^{(p)}, \mathbf{h}_i^{(p)} \right\rangle + \int_0^\infty \left(A_p(\mathbf{k}') + \bar{A}_p(\mathbf{k}')\right)\left\langle \mathbf{e}_j^{(p)}, \mathbf{H}_i^{(p)}(\mathbf{k}') \right\rangle dk_z' \\
&+ \int_0^\infty \left(A_s(\mathbf{k}') + \bar{A}_s(\mathbf{k}')\right)\left\langle \mathbf{e}_j^{(p)}, \mathbf{H}_i^{(s)}(\mathbf{k}') \right\rangle dk_z' \\
B_p(\mathbf{k}) + \bar{B}_p(\mathbf{k}) &= (a_p + \bar{a}_p)\left\langle \mathbf{E}_j^{(p)}(\mathbf{k}), \mathbf{h}_i^{(p)} \right\rangle + \int_0^\infty \left(A_p(\mathbf{k}') + \bar{A}_p(\mathbf{k}')\right)\left\langle \mathbf{E}_j^{(p)}(\mathbf{k}), \mathbf{H}_i^{(p)}(\mathbf{k}') \right\rangle dk_z' \\
&+ \int_0^\infty \left(A_s(\mathbf{k}') + \bar{A}_s(\mathbf{k}')\right)\left\langle \mathbf{E}_j^{(p)}(\mathbf{k}), \mathbf{H}_i^{(s)}(\mathbf{k}') \right\rangle dk_z' \\
B_s(\mathbf{k}) + \bar{B}_s(\mathbf{k}) &= \int_0^\infty \left(A_s(\mathbf{k}') + \bar{A}_s(\mathbf{k}')\right)\left\langle \mathbf{E}_j^{(s)}(\mathbf{k}), \mathbf{H}_i^{(s)}(\mathbf{k}') \right\rangle dk_z' \qquad (34)
\end{aligned}
$$

Implementing the same numerical method as earlier in this paper, two $2N+1 \times 2N+1$ coupling matrices specifying a $4N+2 \times 4N+2$ scattering matrix completely specifies the 3D problem. The discussion of results of a numerical implementation of the expressions above is beyond the scope of the current publication, however, the approach is still semi-analytical in nature. Although not presented here, it is noteworthy to add that in the case of small incidence angles, coupling to TE waves can be neglected yet maintain good field matching of the TM components provided H_x and E_y are small.

V. CONCLUSIONS

We report an efficient, robust and exact semi-analytical mode matching approach for modeling the problem of electromagnetic surface wave scattering at single and multiple surface discontinuities. Excellent agreement between this method and a numerical finite element method was demonstrated for the 2D problem. However, the semi-analytical method solved the field profiles significantly faster (> 25 times for typical 2D problems) and provided a wealth of other useful information such as modal coupling and scattering coefficients that clearly highlight its applicability for the efficient and intuitive design of plasmonic devices

based on structural non-uniformities including nano-plasmonic cavity structures.

Reflectivity, transmission and scattering of surface plasmons impinging at normal incidence onto single and multiple dielectric and metallic surface discontinuities were analyzed in detail. In the cases studied, the reflectivities for low to high surface plasmon wavevectors followed very closely the Fresnel reflection coefficient for bulk waves, while the transmission and radiation scattering share the remaining power. In the case of dielectric discontinuities, small deviations from the expected Fresnel values occurred for high to low surface plasmon wavevectors due to increased scattering; this effect was much larger in the case of metallic discontinuities. The modal analysis provided an explanation for this behavior: in the case of low to high wavevector surface plasmons, very little back-scattered radiation indicated good field matching in the incident medium. On the hand, for high to low wavevector surface plasmons, significantly larger back-scattered radiation indicated poor mode matching in the incident medium and therefore a deviation from the effective Fresnel picture. Further analysis of the radiated fields shows that the majority of the scattered power is in the forward direction.

The open guiding geometry (i.e. with dielectric half spaces) which readily supports radiating waves allows only moderate DBR reflection coefficients of 80 - 90 %. The physical reason lies in the competition between intrinsic and scattering loss: on the one hand, the reflectivity (DBR penetration) of the guided mode increases (decreases) with increasing "strength" (ratio of the guided mode's effective indices) of the non-uniformities thereby reducing propagation losses; on the other hand, increasing strength of the non-uniformity results in increasing the scattering into radiation modes. The competition between these two mechanisms for the total conserved energy results in an optimal (maximum) but non-unity reflection coefficient. Since the main source of loss is through modal scattering, alternative guiding geometries that suppress radiation modes, such as gap and channel structures, would be superior for the design of strongly reflective DBRs.

We also discuss the method's adaptation to the 3D situation: with a moderate increase in calculation complexity, a semi-analytical solution is still possible. This is the first suggestion in the literature that the 3D problem is analytically accessible. This is particularly important since the accuracy of finite-difference and finite-element approaches is typically compromised in three-dimensional calculations as a result of coarse spatial discretization and insufficient computational volumes imposed by drastic memory requirements.

VI. APPENDIX: DERIVATION OF $I(\mathbf{k}_i, \mathbf{k}_j)$

The integral from Eqn. (13) is first written in terms of the field functions for $z < 0$ in the metal region and for $z \geq 0$ in the dielectric region as follows:

$$\begin{aligned} I(\mathbf{k}_i, \mathbf{k}_j) &= \int_{-\infty}^{\infty} (1 - H(z)) \frac{\psi_{i,-}(z)\psi_{j,-}(z)}{\varepsilon_{m,i}} dz + \int_{-\infty}^{\infty} H(z) \frac{\psi_{i,-}(z)\psi_{j,-}(z)}{\varepsilon_{d,i}} dz \\ &= (1 - r_i)(1 - r_j) \int_{-\infty}^{\infty} (1 - H(z)) \frac{\exp\left(-\mathrm{i}(k_{zm,i} + k_{zm,j})z\right)}{\varepsilon_{m,i}} dz \\ &+ \int_{-\infty}^{\infty} H(z) \frac{\exp\left(-\mathrm{i}(k_{zd,i} + k_{zd,j})z\right)}{\varepsilon_{d,i}} dz - r_i \int_{-\infty}^{\infty} H(z) \frac{\exp\left(\mathrm{i}(k_{zd,i} - k_{zd,j})z\right)}{\varepsilon_{d,i}} dz \\ &- r_j \int_{-\infty}^{\infty} H(z) \frac{\exp\left(-\mathrm{i}(k_{zd,i} - k_{zd,j})z\right)}{\varepsilon_{d,i}} dz + r_i r_j \int_{-\infty}^{\infty} H(z) \frac{\exp\left(\mathrm{i}(k_{zd,i} + k_{zd,j})z\right)}{\varepsilon_{d,i}} dz \end{aligned} \tag{35}$$

Each of the integrals in this expression involve the Fourier transform of the Heaviside step function, $H(z)$, where, $H(z < 0) = 0$, $H(z = 0) = 0.5$, $H(z > 0) = 1$ and

$$\int_{-\infty}^{\infty} H(z) \exp\left(\pm \mathrm{i}kz\right) dz = \pi\delta(k) \mp \frac{\mathrm{i}}{k} \tag{36}$$

Substitution into Eqn. (35) gives the general expression for the coupling coefficients integral,

$$\begin{aligned} I(\mathbf{k}_i, \mathbf{k}_j) &= \pi \frac{(1 - r_i)(1 - r_j)}{\varepsilon_{m,i}} \delta(k_{zm,i} + k_{zm,j}) - \mathrm{i} \frac{(1 - r_i)(1 - r_j)}{\varepsilon_{m,i}(k_{zm,i} + k_{zm,j})} \\ &+ \pi \frac{(1 + r_i r_j)}{\varepsilon_{d,i}} \delta(k_{zd,i} + k_{zd,j}) + \mathrm{i} \frac{(1 - r_i r_j)}{\varepsilon_{d,i}(k_{zd,i} + k_{zd,j})} \\ &- \pi \frac{(r_i + r_j)}{\varepsilon_{d,i}} \delta(k_{zd,i} - k_{zd,j}) + \mathrm{i} \frac{(r_i - r_j)}{\varepsilon_{d,i}(k_{zd,i} - k_{zd,j})} \end{aligned} \tag{37}$$

Here, note that only one delta function is retained, since the signs of the $\mathbf{k}$ vectors have already been accounted. Therefore, Eqn. (37) simplifies to,

$$\begin{aligned} I(\mathbf{k}_i, \mathbf{k}_j) &= -\pi \frac{(r_i + r_j)}{\varepsilon_{d,i}} \delta(k_{zd,i} - k_{zd,j}) + \mathrm{i} \frac{(1 - r_i r_j)}{\varepsilon_{d,i}(k_{zd,i} + k_{zd,j})} \\ &+ \mathrm{i} \frac{(r_i - r_j)}{\varepsilon_{d,i}(k_{zd,i} - k_{zd,j})} - \mathrm{i} \frac{(1 - r_i)(1 - r_j)}{\varepsilon_{m,i}(k_{zm,i} + k_{zm,j})} \end{aligned} \tag{38}$$

This expression reduces to the form shown in Eqn. (13) by factorizing $(1 - r_i)(1 - r_j)$ and noting that,

$$\frac{1 + r_i}{1 - r_i} = \frac{\varepsilon_{m,i} k_{zd,i}}{\varepsilon_{d,i} k_{zm,i}} \tag{39}$$

Acknowledgments

This work was supported by the U.S. Air Force Office of Scientific Research MURI program under Grant No. FA9550-04-1-0434 and by the NSF Nanoscale Science and Engineering Center (NSEC) under Grant No. DMI-0327077.

---

* Electronic address: r.oulton@berkeley.edu

[1] H. Raether, *Surface Plasmons on Smooth and Rough Surface and on Gratings* (Springer-Verlag, Berlin, 1986)

[2] S. A. Maier, P. G. Kik, H. A. Atwater, S. Meltzer, E. Harel, B. E. Koeland and A. A. G. Requicha, Nature Materials **2**, 229 (2003)

[3] P. Berini, Optics Lett. **24**, pp 1011-1013 (1999)

[4] S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, J.-Y. Laluet, T. W. Ebbesen, Nature, **440**, pp. 508-511 (2006)

[5] K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. R. Dasari, and M.S. Feld, Phys. Rev. Lett. **78**, pp. 1667-1670 (1997)

[6] K-H Su, S. Durant, J. M. Steele, Y. Xiong, C. Sun and X. Zhang, J. Chem. Phys. **100**, pp 3964-3968 (2006)

[7] N. Fang, H. Lee, C. Sun and X. Zhang, Science **308**, pp 534-537 (2005)

[8] H. T. Miyazaki and Y. Kurokawa, Phys. Rev. Lett **96**, 097401 (2006)

[9] D. J. Bergman and M. I. Stockman, Phys. Rev. Lett. **90**, 027402 (2003)

[10] J. R. Krenn, H. Ditlbacher, G. Schider, A. Hohenau, A. Leitner and F. R. Aussenegg, Journal of Microscopy **209**, 3, pp. 167 - 172 (2003)

[11] A. V. Zayats, I. I. Smolaninov and A. A. Maradudin, Physics Reports **408**, pp. 131-314 (2005)

[12] B. Wang and G. P. Wang, App. Phys. Lett. **87**, pp 013107 (2005)

[13] L. Zhou, X-Q. Yu and Y-Y Zhu, App. Phys. Lett. **89**, pp. 051901 (2006)

[14] D. F. P. Pile and G. K. Gramotnev, Appl. Phys. Lett. **86**, 161101 (2005)

[15] P. J. B. Clarricoats and K. R. Slinn, Proc. IEEE **114**, 7, pp.878-886 (1967)

[16] F. Pincemin, A. A. Mardudin, A. D. Boardman and J. -J. Greffet, Phys. Rev. B **50**, 15261 (1994)

[17] T. Sondergaard and S. I. Bozhevolnyi, Phys. Rev. B **69**, 165405 (2003)

[18] B. T. Draine, Astrophys. J. **333**, 848 (1988)

[19] A. B. Evlukhin, S. I. Bozhevolnyi, A. L. Stepanov and J. R. Krenn, Appl. Phys. B **84**, 29 (2006)

[20] V. V. Schevchenko, Golem, Bolder Colorado (1971).

[21] S. F. Mahmoud and J. C. Beal, IEEE MTT **23**, 2, pp 193-198 (1975)

[22] G. I. Stegeman, A. A. Maradudin and T. S. Rahman, Phys. Rev. B **23**, 2576 (1981)

[23] A. I. Voronko, L. G. Klimova and G. N. Shkerdin, Solid State Commun. **61**, 6, pp. 361-364 (1987)

[24] H. A. Jamid and S. J. Al-Bader, IEEE Photon. Technol. Lett. **9**, 2, pp 220 - 222 (1997)

[25] R. E. Collin, *Field Teory of Guided Waves*, 2nd Edition (Wiley/IEEE Press, New York, 1991)

[26] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, New York, 1985).

[27] P. B. Johnson and R. W. Christie, Phys. Rev. B **6**, pp. 4370-4379 (1972)

[28] E. N. Economou, Phys. Rev. **182**, pp 539-554 (1969)

[29] J. A. Dionne, L. A. Sweatlock, H. A. Atwater and A. Polman, Phys. Rev. B **73**, 035407 (2006)