Enhanced transmission with coaxial nanoapertures: Role of cylindrical surface plasmons

Michael I. Haftel,1 Carl Schlockermann,2 and G. Blumberg3
1Center for Computational Materials Science, Naval Research Laboratory, Washington, DC 20735-5343, USA
2Munich University of Applied Sciences, Munich, Germany, and
3Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

(Received 8 August 2006; revised manuscript received 11 October 2006; published 4 December 2006)

We simulate the optical fields and optical transmission through nanoarrays of silica rings embedded in thin gold films using the finite-difference-time-domain method. By examining the optical transmission spectra for varying ring geometries we uncover large enhancements in the transmission at wavelengths much longer than the usual cutoffs for cylindrical apertures or where the usual planar surface plasmons or other periodic effects from the array could play a role. We attribute these enhancements to closely coupled cylindrical surface plasmons on the inner and outer surfaces of the rings, and this coupling is more efficient as the inner and outer ring radii approach each other. We confirm this hypothesis by comparing the transmission peaks of the simulation with cylindrical surface plasmon (CSP) dispersion curves calculated for the geometries of interest. One important result is that a transmission peak appears in the simulations close to the frequency where the longitudinal wave number $k_z$ in the ring satisfies $k_z = m \pi / L$, where $m$ is an integer and $L$ the length of the aperture, for a normal CSP TE$_1$ or TM$_1$ mode. The behavior of the CSP dispersion is such that propagating modes can be sent through the rings for ever longer wavelengths as the ring radii approach, whereas the transmission decreases only in proportion to the ring area.

DOI: 10.1103/PhysRevB.74.235405

PACS number(s): 78.67.-n, 73.20.Mf, 42.79.Ag

I. INTRODUCTION

The phenomenon of enhanced optical transmission (EOT) through nanoarrays of apertures has attracted a great deal of attention ever since the publication of Ebbesen et al.’s work in 1998.1 This work, while having far-reaching implications for optical device development (as well as for devices in the RF and microwave regime), has generated a lot of controversy concerning the origin of the enhanced transmission. While originally thought to be attributable to surface plasmons (SP’s)1–3 on the upper and lower surfaces of the metal film in which the apertures were embedded, other investigators have shown that other effects including shape resonances4,5 and resonant coupling between the elements of the array,6,7 present even for perfect conductors where SP’s are not present,6,7 and other diffractive effects from evanescent waves,8 can account for EOT. The explicit role of SP’s has not been completely resolved since the other enhancements often occur at frequencies very close to those where surface plasmons would be launched.

Apart from the role of SP’s, EOT is very sensitive to the shape of apertures. For example, the aspect ratio of rectangular apertures can have a large influence on EOT.4 In a series of papers8–12 considering arrays of coaxial ring (CR) apertures, Baida and co-workers have shown in finite-difference-time-domain (FDTD) simulations that the EOT from coaxial ring apertures can be much larger than from cylindrical apertures. Furthermore, the transmission peaks for coaxial gold rings are considerably redshifted from those of either perfectly conducting coaxial rings or cylindrical holes. Such coaxial structures have recently been fabricated by Salvi et al.,13 who find good agreement between their measured transmission intensities and FDTD simulations, at least for wavelengths up to the maximum measured of 900 nm. These findings demonstrate that one can exceed the diffraction limit to a much greater degree when these CR structures are used compared to apertures with a cylindrical shape or other open structures. This could have far-reaching consequences for optical devices in sensing, detection, and communication applications. The main purpose of this paper is to analyze the underlying physics of the coaxial geometry and to show that cylindrical surface plasmon (CSP) resonances are primarily responsible for this behavior. We have previously sketched the role of these resonances for such CR structures.14 The present paper gives a detailed analysis of their optical properties.

The previous investigators attribute the enhancements at long wavelengths to the mode structure of the individual coaxial rings (CR),10,12,13 presumably a TE$_1$ mode, and not to the periodic structure. How the behavior of these modes is determined by the properties of the gold film (and gold core of the rings) has not been delineated. Some of the normal modes for a CR in a perfect conductor do differ considerably from those of a cylindrical aperture.12,15 For example, the lowest frequency TE$_1$ mode has a cutoff $\approx \pi (R_1 + R_2)$,12,15 where $R_1$ and $R_2$ are the inner and outer ring radii. This gives the somewhat nonintuitive property that, for a constant outer radius $R_2$, the cutoff approaches a constant value $2 \pi R_2$ as the ring closes ($R_1 \rightarrow R_2$). (Of course, the transmission vanishes in this limit as the ring closes up.) Also, the CR supports a TEM$_0$ mode without cutoff. This mode, however, has radial symmetry and cannot be readily excited by an incoming plane wave of linear or circular polarization, and thus would not be seen in a FDTD simulation or typical experiments. In this work we show that the cutoff wavelength increases indefinitely in the above limit, and this is a consequence of the negative dielectric constant of metals and the ensuing cylindrical surface plasmons.

The above work on CR structures9–12 motivated us to carry out FDTD simulations to guide possible device fabrication processing. These simulations strongly suggest that the source of the unusual EOT for the coaxial ring apertures...
are cylindrical surface plasmon (CSP) resonances, with the propagating CSP’s on each of the metal-dielectric interfaces becoming strongly coupled as the ring narrows. In the remainder of this paper we present the evidence for such an interpretation. Section II briefly describes our model system and the basics of our FDTD simulation, including the parameters of an extended Drude model for the dielectric constant of gold (Au). Section III examines the analytic features of the normal mode problem, as a function of the inner and outer ring ratio, for a coaxial cylinder with a metal core and a dielectric ring embedded in a metal film. By diagonalizing the appropriate $8 \times 8$ matrix we derive dispersion curves for CSP normal modes and from these derive the resonant conditions, which allow us to predict the peak positions as a function of ring geometry. We show that the cutoff for propagating modes increases indefinitely as the ring narrows with a resultant redshifting of the transmission peaks. We assess the role of losses in limiting this behavior. In Sec. IV we carry out FDTD simulations for different CR geometries. These simulations confirm the analytic trends established in Sec. III. While the transmission peaks increasingly redshift with narrower rings, the intensity of transmission decreases no faster than the exposed ring area, and the transmitted energy is two to four times that incident on the ring. In Sec. V we summarize our results and present conclusions.

II. MODEL SYSTEM

Our “basic” model consists of an infinite array of coaxial silica (index of refraction=1.46) rings spaced in a square pattern at $d=555$ nm intervals, with inner and outer radii $R_1=50$ nm and $R_2=100$ nm, imbedded in an $L=290$ nm thick Au film. The metal film is itself imbedded in a uniform silica medium. The incident light is a linearly polarized plane wave traveling in the silica medium incident normally on the Au film (from the bottom). Figure 1 illustrates our model system. We consider variations of the CR geometry, namely changes in $R_1$, $L$, and $d$, as well as considering our basic geometry with Au replaced by a perfect conductor (PC).

We simulate the fields with the NRL High Accuracy Scattering and Propagation (HASP) code. This code employs a nonstandard finite difference (NSFD) algorithm that yields much smaller amplitude and dispersion errors than the standard Yee algorithm when used on a coarse ($\lambda/10$) grid. We employ periodic boundary conditions in the $x$ and $y$ directions (parallel to the film) and Mur absorbing boundary conditions in $z$. The spatial step is 6.9375 nm in all directions, and the typical time step is 0.02 fs. The incident pulse is a plane wave of period 3.0 fs ($\lambda=900$ nm in vacuum, 623 nm in silica) with a Gaussian envelope with a full width half maximum (FWHM) of 0.809 fs. The frequency dependence of the fields at any point is calculated from the simulation using standard fast Fourier transform (FFT) techniques. We compute the transmission intensity at a given frequency by comparing the real part of the $z$ component of the Poynting vector integrated over a plane 700 nm above the film with the corresponding quantity in the incident beam. (This intensity is stable as the height above the film is varied from 300–1000 nm).

We have verified the HASP code a number of different ways. Simulations of Mie scattering from cylinders and spheres reproduce the analytic results to about 2% even on a coarse ($\lambda/100$) grid. Here we are using about a $\lambda/100$ grid. The HASP code also gives good agreement with the near-surface fields in the surface plasmon jet observations of Egorov et al. HASP also reproduces the positions of the transmission peaks observed by Martin-Moreno et al. for a free standing silver film with cylindrical apertures (although the magnitude of the peaks differs from the experiment).

We model the dielectric constant of Au with an extended Drude model, i.e.,

$$\varepsilon_0(\omega) = \varepsilon_1 - i\varepsilon_2/\omega\tau(1 - i\omega\tau) + 4\pi\sigma/\omega,$$  

(1)

where $\varepsilon_1$, $\varepsilon_2$, $\sigma$, and $\tau$ are fit to the experimental dielectric constant. This form is a generalization from the Drude model that allows us to fit the dielectric constant over a much larger range of wavelengths than the pure Drude model. Thus we are not required to carry out numerous simulations over limited ranges of wavelengths to cover the full range of interest. Values of $\varepsilon_1=7.919$, $\sigma'=0.0056$ fs$^{-1}$ (we use Gaussian units), $\varepsilon_2=-14.395.1016$ and $\tau=9.00$ fs yield a good fit to the experimental data of Johnson and Christy for $\lambda =600–2000$ nm. Figure 2 illustrates this fit for $\lambda >450$ nm.

The real part fits the experimental data well for $\lambda >520$ nm, while the imaginary part fits well for $\lambda >630$ nm. We include more recent experimental data of Lynch and Hunter up to 10 000 nm, and our fit agrees reasonably well at least up to
Our simulation results closely resemble those of Baida et al.\textsuperscript{10} as does the conclusion that some of the peaks are not due to planar surface plasmons. Allowing for the slightly different periodicity (theirs is 600 nm) and dielectric medium (they use glass, $\varepsilon=2.34$) the peak positions and features of our Fig. 3 for $R_1=0$ are very similar to those of Fig. 4 of Ref. 10, and the features in Fig. 3 for $R_1=50$ nm are very similar to those of Fig. 8 of Ref. 10. For example, for $R_1=50$ nm and $R_2=100$ nm (see Fig. 8, Ref. 10), both our results and those of Baida et al.\textsuperscript{10} exhibit a pair of peaks (at 1000 nm and 903 nm in Fig. 3) at long wavelengths separated by $\sim 100$ nm, followed by a triad of peaks at wavelengths 100–200 nm shorter (at 654, 714, and 763 nm in Fig. 3), and a rather small “bump” at a somewhat shorter wavelength (598 nm in Fig. 3, at about 680 nm in Fig. 8, Ref. 10). An even smaller bump $\sim 40$ nm blueshifted from the latter that is hard to discern in either figure. The relative magnitudes of these features in our simulations and in Ref. 10 are comparable in magnitude. In addition to the aforementioned SP modes, corresponding Wood’s anomalies (WA), in our case theoretically at 810 and 572 nm, could play a role in accounting for some of these simulated features. The higher order features at the shorter wavelengths ($<700$ nm) in our simulations and those in Ref. 10 do not show up as clearly as in some other simulations, such as those of Chang et al.\textsuperscript{27} who consider 200 nm open square apertures. Here the decreased exposed area of our coaxial ring apertures (compared, e.g., to open circular or square apertures) may partially explain the repression of these features. Indeed, these higher order features are further repressed in the $R_1=75$ nm and $R_1=85$ nm results in Fig. 3. There are some peaks, however, e.g., the peaks at 1000 nm and 763 nm in Fig. 3 (with corresponding peaks at 1150 and 900 nm in Fig. 8 of Ref. 10) that cannot be explained by PSP or WA effects.

We have also simulated the system studied by Salvi et al.\textsuperscript{13} We obtain peaks at 660 nm and 1300 nm versus their results of 700 and 1320 nm. The difference is likely due to the different treatment of the dielectric constant of Au. (They use different pure Drude model fits over limited wavelength intervals.)

How do the modes of the individual coaxial rings account for the behavior of the transmission spectra shown in Fig. 3? To investigate this question we first carry out a normal mode analysis of the geometry of interest (Fig. 1) in cylindrical coordinates to obtain CSP dispersion curves for a single (infinitely long) CR completely surrounded by Au. We seek the relation between the longitudinal wave number $k_z$ in the CR, the azimuthal state $n$, and the frequency $\omega$, denoted by $a_n(k_z)$ or the corresponding wavelength $\lambda_n(k_z)$. We follow the procedure described by Schröter and Dereux,\textsuperscript{28} who derive the CSP dispersion curves for a metal ring embedded in a dielectric. Kushawa et al.\textsuperscript{29,30} have developed Green’s function techniques for calculating dispersion curves for coaxial\textsuperscript{29} and multiaxial\textsuperscript{30} structures with arbitrary dielectric constants with applications to quantum wire and nanotubelike structures.

We straightforwardly adapt the methods of Ref. 27, except now with a dielectric ring embedded in a metal. We concentrate on $n=1$ as this is the azimuthal mode most readily excited by a linearly polarized plane wave. A normal mode exists whenever the appropriate boundary conditions on $E_{\phi}, E_{\rho}, H_{\phi}$, and $H_{\rho}$, at the two cylindrical metal-dielectric interfaces can be satisfied, which amounts to finding the zeroes of the determinant of an $8 \times 8$ matrix involving TE and TM coefficients for the various cylindrical Bessel (or Neumann or Hankel) functions in the radial wave function for the three radial regions. (For $n=0$ or $k_z=0$ pure TE or TM modes can be found; otherwise, the CSP modes are TE-TM combinations). We will show that these modes have a surface
plasmon-like character, i.e., the fields are concentrated near the metal-dielectric cylindrical interfaces.

In general, all the components of \( \mathbf{E} \) and \( \mathbf{H} \) can be expressed in terms of the TM and TE wave functions, \( \Psi_{\text{TM}} \) and \( \Psi_{\text{TE}} \), and their first and second derivatives with respect to the cylindrical coordinates \( \rho, \phi, \) and \( z \), where, in a radial region \( i \) for azimuthal mode \( n \),

\[
\Psi_{\text{TM}} = \sum_{m} A_{mn} F_{n}^{m}(k_{m} \rho) \exp(ik_{z}z) \exp(in\phi),
\]

\[
\Psi_{\text{TE}} = \sum_{m} B_{mn} F_{n}^{m}(k_{m} \rho) \exp(ik_{z}z) \exp(in\phi),
\]

where \( F_{n}^{m}(z) \) is a cylindrical Bessel-type function whose type (Bessel, Neumann, or Hankel) is labeled by \( m \), and \( k_{i} \) is the radial wave number that satisfies

\[
\begin{align*}
\begin{array}{cccccccc}
J_{1} & -J_{2} & -N_{2} & 0 & 0 & 0 & 0 & 0 \\
-\alpha \omega J_{1} / k_{1} c & -\alpha \omega J_{2} / k_{2} c & -\alpha \omega N_{2} / k_{2} c & 0 & 0 & 0 & 0 & 0 \\
0 & -J_{3} & -N_{3} & H_{4} & 0 & 0 & 0 & 0 \\
0 & -i \alpha \omega J_{1} / k_{1} c & -i \alpha \omega N_{2} / k_{2} c & i \alpha \omega H_{4} / k_{3} c & 0 & 0 & 0 & 0 \\
-nk_{1} J_{1} / k_{1} c^{2} R_{1} & nk_{2} J_{2} / k_{2} c^{2} R_{1} & nk_{2} N_{2} / k_{2} c^{2} R_{1} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -J_{3} & -N_{3} & H_{4} & 0 \\
0 & nk_{1} J_{1} / k_{1} c^{2} R_{2} & nk_{2} J_{2} / k_{2} c^{2} R_{2} & nk_{2} N_{2} / k_{2} c^{2} R_{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & i \alpha \omega J_{1} / k_{1} c & i \alpha \omega J_{2} / k_{2} c & i \alpha \omega N_{2} / k_{2} c & i \alpha \omega H_{4} / k_{3} c,
\end{array}
\end{align*}
\]

where \( J_{i} = J_{i}(k_{i} R_{i}) \), \( J_{2} = J_{n}(k_{2} R_{2}) \), \( J_{3} = J_{n}(k_{2} R_{3}) \), with similar labeling for the \( N_{n} \) and \( H_{4} = H_{4}^{1}(k_{3} R_{3}) \). In our case, where the core of the coaxial rings and the film are both Au, \( \varepsilon_{1} = \varepsilon_{2} = \varepsilon_{\text{Au}}(\omega) \), and \( k_{1} = k_{2} \). (We have assumed that \( \mu_{1} = \mu_{2} = \mu_{3} = 1 \).) The continuity conditions are summarized in the matrix equation \( \mathbf{M} \mathbf{z} = 0 \), where \( \mathbf{a} \) is the coefficient vector whose components are \( (A_{1,1}, A_{2,1}, A_{3,1}, B_{1,1}, B_{2,1}, B_{3,1}) \) of Eq. (2), and the eight components of the vector \( \mathbf{Ma} \) when equated to zero, give the continuity conditions on \( E_{r}(R_{i}), H_{r}(R_{i}), E_{z}(R_{i}), H_{z}(R_{i}), H_{n}(R_{1}), H_{n}(R_{2}), H_{n}(R_{3}), \) and \( E_{\phi}(R_{i}) \), respectively. If the metal is perfectly conducting, then pure TE and TM solutions are possible. In the case of a perfectly conducting metal with \( R_{1} > 0 \), a cylindrically symmetric TEM mode is also possible, and this mode has no cutoff. For a real \( (\epsilon \) finite) pure TE or TM solutions are coupled unless \( n = 0 \) or \( k_{z} = 0 \) [Eq. (4)], and except for these cases the solutions are admixtures of TE and TM. The TEM mode is still possible, but the cylindrical symmetry of this mode means that it cannot be excited by an incident linearly (or circularly) polarized plane wave. The evidence to date indicates that the long wavelength modes are predominantly TE, which is most easily excited by incident plane waves.

Figure 4 gives the \( n=1 \) dispersion curves for Au, expressed in terms of the wavelength \( \lambda_{n}(k_{z}) \) for six values of the inner radius \( R_{1} \) (\( R_{2}=100 \) nm). To simplify matters we have used only the real part of the dielectric constant of Au. Including the imaginary part hardly alters these curves for the real part of \( \lambda_{n}(k_{z}) \), but of course a finite width will be introduced (usually widening to \( 10-30 \) nm in \( \lambda \)). The vertical lines correspond to values of \( k_{z} \) that satisfy the relation \( k_{z} L = m \pi \) for various \( L \). The condition

\[
k_{z} L = m \pi
\]

is the general condition for a longitudinal standing wave in the cylinder, the intersection of any of these lines (the vertical axis) with a given curve will give the wavelength for \( m=1 \) (\( m=0 \)) at which the cylindrical mode is in resonance with the longitudinal mode. [Eq. (5)] is the usual longitudinal resonance condition for a cavity with perfectly conducting walls; we assume it is approximately true for the real metal, which our later simulations bear out.] Physically, we tentatively identify these as CSP resonances, as our following analysis will confirm that these solutions have the characteristics of surface plasmons propagating on the metal-dielectric interfaces because of the negative dielectric constant of the metal. Thus a picture of CSP assisted extraordinary transmission takes form. As previously noted, the EOT is primarily driven by TE, wave guide modes \( (k_{z} \gg 0) \), which differs from the EOT produced by planar surface plasmons, where the
FIG. 4. The \( n=1 \) CSP dispersion curves for silica CR’s embedded in a gold film for the basic case but with various \( R_1 \) values. The solid vertical line indicates the values of \( k_z \) corresponding to \( k_z = \pi/L \) for \( L=290 \) nm, while the dotted vertical lines indicate corresponding values of \( k_z \) for \( L=420 \) and 1000 nm.

latter involves evanescent modes propagating through the aperture.

We find a very close correspondence between the resonant frequencies in the simulations (Fig. 3) with those obtained with the dispersion curves with the resonance condition of Eq. (5) (see Fig. 4). For example, for \( n=1 \) with our basic model \((R_1=50 \) nm, \( L=290 \) nm) the dispersion analysis predicts \( m=0 \), 1 peaks at 979 nm and 789 nm, respectively, while the simulation gives these peaks at 1000 nm and 757 nm. For \( R_1=75 \) nm the dispersion analysis gives 1376 nm and 1035 nm, while the simulations in Fig. 3 give 1406 nm and 1110 nm. The simulation results closely follow the trends of the dispersion analysis, but tend to be redshifted somewhat from the analytical results. Differences can come from several sources including losses, presence of other modes, interference of CSP and PSP modes, and the “staircase” representation of the CR on a finite rectangular mesh, as well as the condition of Eq. (5) being approximate.

Figures 5(a) and 5(b) show the radial dependence of the dominant field components for a typical solution corresponding to \( k_z = 0.0108 \) nm\(^{-1} \) for inner radius equal to 50 nm and 85 nm, respectively. The magnitude of the fields is maximum at the inner metal-dielectric interface with cusplike curvature at both interfaces like that of decaying exponentials away from the interfaces on either side. (For some components the field is actually decreasing or nearly constant as it approaches the outer radius, but with positive curvature.) This is the expected behavior of a cylindrical surface plasmon, i.e., surface states on the inner and outer radius of the ring. As the ring narrows (see the 85 nm case in Fig. 5) the fields do not have a chance to substantially fall off away from the inner radius, and the average field intensity in the dielectric ring is higher than for a wider ring. In a sense the CSP wave functions from the two interfaces are overlapping more. Another possible physical picture is that the surface plasmons from the two surfaces are more coherent, i.e., that while the surface plasmons have the same angular momentum (the same azimuthal quantum number \( n \)) they also approach having the same angular velocity as \( R_1 \rightarrow R_2 \).

From the dispersion curves in Fig. 4 it is evident that the wavelengths at the peaks for \( m=0 \) and \( m=1 \) increase rapidly as \( R_1 \rightarrow R_2 \). What is the limit for such behavior, i.e., does the peak redshift indefinitely? To determine this dependence on the width of the ring we consider the condition for \( \text{det}(M) = 0 \) in the limit \( \Delta R = R_2 - R_1 \rightarrow 0 \). For \( n=1 \) and \( k_z=0 \) the TE and TM modes decouple and the TE modes are determined by a \( 4 \times 4 \) matrix

\[
M = \begin{pmatrix}
J_1(k_m R_1) & -J_1(k_d R_1) & -N_1(k_d R_1) & 0 \\
\frac{i \omega J_1'(k_m R_1)}{k_d c} & \frac{i \omega J_1'(k_d R_1)}{k_d c} & \frac{i \omega N_1'(k_d R_1)}{k_d c} & 0 \\
0 & -J_1(k_d R_2) & -N_1(k_d R_2) & H_1^{(1)}(k_m R_2) \\
0 & \frac{i \omega J_1'(k_d R_2)}{k_d c} & \frac{i \omega N_1'(k_d R_2)}{k_d c} & \frac{i \omega H_1^{(1)}(k_m R_2)}{k_m c}
\end{pmatrix}
\]

(6)

The determinant of \( M \) is straightforward to evaluate. Since, for \( \lambda > 600 \) nm \( |\varepsilon_m| >> |\varepsilon_d| \), thus \( k_m \gg k_d \), we express the determinant in powers of \( k_m/k_d \).
The last square bracket in Eq. (7) is of first order and reduces to
\[ \left[ N_1'(k_m R_1) J_1'(k_m R_1) - J_1'(k_m R_1) N_1'(k_m R_1) \right] k_d \Delta R, \]
which can be further reduced, by application of the Bessel function differential equation
\[ w''(z) = -w'(z)/z - (1 - 1/z^2)w(z), \]
to
\[ \left( 1 - 1/(k_d^2 R_1^2) \right) \left[ N_1'(k_m R_1) J_1'(k_m R_1) - J_1'(k_m R_1) N_1'(k_m R_1) \right] k_d \Delta R. \]
The first order contribution to the first square bracket in the \( k_m/k_d \) term in of Eq. (7) vanishes leaving only a zero order contribution. The second square bracket in the \( k_m/k_d \) term becomes, upon expansion to first order
\[ J_1(k_m R_1) N_1'(k_m R_1) - N_1(k_m R_1) J_1'(k_m R_1) + k_d \Delta R \left[ J_1(k_m R_1) N_1'(k_m R_1) - N_1(k_m R_1) J_1'(k_m R_1) \right], \]
which reduces to, upon application of the Bessel function differential equation
\[ J_1(k_m R_1) N_1'(k_m R_1) - N_1(k_m R_1) J_1'(k_m R_1) - \Delta R/R_1 \left[ J_1(k_m R_1) N_1'(k_m R_1) - N_1(k_m R_1) J_1'(k_m R_1) \right]. \]
The last square bracket in Eq. (7) vanishes in zero order and only the first order contribution
\[ \left[ N_1(k_m R_1) J_1'(k_m R_1) - J_1(k_m R_1) N_1'(k_m R_1) \right] k_d \Delta R \]
remains. Upon application of all of these relations, a common factor
\[ W(k_m R_1) = N_1(k_m R_1) J_1'(k_m R_1) - J_1(k_m R_1) N_1'(k_m R_1) \]
appears in Eq. (7). Equation (7) now becomes
\[
\begin{align*}
\text{det}(M) = \frac{\omega^2}{k_m^2 c^2} & \left\{ \frac{k_m^2}{k_d^2} J_1(k_m R_1) H_1(k_m R_2) \left[ 1 - \frac{1}{k_d^2 R_1^2} \right] k_d \Delta R \\
- \frac{k_m}{k_d} \left[ J_1(k_m R_1) H_1'(k_m R_2) - J_1'(k_m R_1) H_1(k_m R_2) \right] \left[ 1 - \frac{1}{R_1} \right] \\
+ J_1'(k_m R_1) H_1'(k_m R_2) k_d \Delta R \right\}.
\end{align*}
\]
Further simplification occurs because of the identity \( W(z) = -2/\pi z \). Further assuming that \( k_m \Delta R \ll 1 \) (to be justified latter), Eq. (10) becomes, ignoring the prefactor since we are only interested in the zero of the determinant,
\[
\text{det}(M) \sim \frac{k_m^2}{k_d^2} J_1(k_m R_1) H_1(k_m R_2) \left[ 1 - \frac{1}{k_d^2 R_1^2} \right] k_d \Delta R,
\]
and the differential equation for \( H_1(z) \) to eliminate \( H_1'(k_m R_1) \) and simplify Eq. (11) further (to first order in \( \Delta R \)),
\[
- J_1(k_m R_1) H_1'(k_m R_1) k_m \Delta R. \]
ENHANCED TRANSMISSION WITH COAXIAL...

\[
\text{det}(M) \sim \frac{k_m^2}{k_d^2} \frac{1}{k_d^2} J_1(k_m R_1) H_1(k_m R_1) \left[ \frac{1}{k_d^2 R_1^2} - 1 \right] k_d \Delta R - \frac{2i}{\pi k_d R_1}
\]

\[
- \frac{k_m}{k_d} J'_1(k_m R_1) H_1(k_m R_1) \Delta R / R_1
\]

\[
+ J'_1(k_m R_1) H'_1(k_m R_1) k_d \Delta R
\]

\[
+ \frac{k_m^2}{k_d^2} \left[ J'_1(k_m R_1) H'_1(k_m R_1) - J_1(k_m R_1) H'_1(k_m R_1) \right]
\]

\[
\times \left[ \frac{1}{k_d^2 R_1^2} - 1 \right] \Delta R + \frac{k_m \Delta R}{k_d R_1} J_1(k_m R_1) H'_1(k_m R_1).
\]

(13)

There are two regimes where our approximations are valid. The first is, for most metals, the optical and near infrared. For these frequencies \(\omega \gg 1\) in the extended Drude model [Eq. (1)], but \(k_d \Delta R \ll 1\), and in this case \(\varepsilon_m = \varepsilon_2 / \omega^2 R^2\) and \(k_m = (\varepsilon_2)^{1/2} / c \tau\). For wavelengths 600–4000 nm \(k_m\) is approximately the constant value 0.04i for Au. The second regime is the extreme \(\omega \rightarrow 0\) limit (which corresponds to the extreme \(\Delta R \rightarrow 0\) limit) where \(\varepsilon_m \sim i / \omega\). With \(\omega\) and \(\Delta R\) as small parameters, in either of the above cases the leading terms proportional to \(\Delta R^0\) and \(\Delta R^2\) are the third and first terms of Eq. (13), respectively. Keeping these leading terms and equating the determinant to zero yields the following condition for a CSP resonance:

\[
k_d^2 = -1/2 i \pi k_m J_1(k_m R_1) H_1(k_m R_1) \Delta R / R_1,
\]

or \(\omega_{res} = c [-1/2 i \pi k_m J_1(k_m R_1) H_1(k_m R_1) \Delta R / R_1]^{1/2} / \varepsilon_d\),

(15)

where \(k_m\) may depend on \(\omega\). In the regime where \(\varepsilon_m\) is real and negative, \(k_m = \kappa_m\), where \(\kappa_m\) is real and positive, \(\omega_{res}\) is real and positive, and can be expressed

\[
\omega_{res} = c [\kappa_m^2 I_1(\kappa_m R_1) K_1(\kappa_m R_1) \Delta R / R_1]^{1/2} / \varepsilon_d,
\]

where \(I_1(z)\) and \(K_1(z)\) are modified Bessel function of the first and second kind.\(^{31}\)

If \(\kappa_m\) is (approximately) constant, the resonant frequency is real, positive, and is decreasing to zero as \(\sim \Delta R^{1/2}\). This behavior of the resonant wavelength, \(\sim \Delta R^{-1/2}\), is roughly consistent with the extraction of resonant frequencies from the dispersion curves (Fig. 4) and simulation results (Fig. 3) in the optical and near infrared regimes. Equation (16) further simplifies if \(\gamma = \kappa_m R_1\) is very large or very small. (In the problem of interest \(\gamma = 4\).) In the large and small \(\kappa_m R_1\) limits

\[
\omega_{res} = c [\kappa_m \Delta R / R_1]^{1/2} / \varepsilon_d,
\]

(17a)

and

\[
\omega_{res} = c \kappa_m [1/2 \Delta R / R_1]^{1/2} / \varepsilon_d,
\]

(17b)

respectively. The fact that the resonant frequency is real and decreases as \(\Delta R^{1/2}\), or that the resonant wavelength increases as \(\Delta R^{-1/2}\), as \(\Delta R \rightarrow 0\), follows from \(\kappa_m\) being real in Eqs. (17a) and (17b), and is a consequence of the negative dielectric constant of the metal. This would not occur for the infinite imaginary dielectric constant of an ideal conductor.

As the ring further narrows, the resonant frequency decreases and eventually we are in the regime where the metal behaves as a good conductor and its dielectric constant becomes purely imaginary, \(\varepsilon_m = 4 \pi i \sigma / \omega\), where in our extended Drude model [Eq. (1)] \(\sigma = \sigma' - i \sigma' / 4 \pi \tau\). In the limit \(\kappa_m R_1 \ll 1\), \(J_1(\kappa_m R_1) H'_1(\kappa_m R_1) \approx -i / \pi\), and we obtain

\[
\omega_{res} = -2 i \pi \sigma \Delta R / \varepsilon_d R_1,
\]

(18)

i.e., the resonant frequency decreases as the ring width, but becomes entirely imaginary. (The frequency in our formulation has a negative imaginary part because we have assumed that the fields have a time dependence \(-\exp(-i \omega t)\), thus a negative imaginary part corresponds to a finite lifetime.) For Au the losses become significant for \(\lambda > 40 \mu m\), whereas the condition \(\kappa_m R_1 \ll 1\) holds for \(\lambda \gg 30 \mu m\). Therefore, as one proceeds from the optical to the THz regime the resonant wavelength increases indefinitely (inversely) as \(\Delta R\), but proceeds off of the real axis, i.e., develops a larger width.

Figure 6 illustrates the dependence of the resonant frequency for \(k_z = 0\) on \(\Delta R\) using the full extended Drude model of Eq. (1), including the imaginary part (i.e., losses). We also include results for the approximation of Eq. (15), which are very close to the exact results for \(\Delta R < 10\) nm. The behavior \(\omega_{res} \sim \Delta R^{1/2}\) is confirmed for \(0.1 \text{ nm} < \Delta R < 10 \text{ nm}\), which roughly corresponds to \(0.1 \text{ fs}^{-1} < \omega < 3.0 \text{ fs}^{-1}\) or wavelengths 0.6 to 180 \(\mu m\). For narrower CR’s (i.e., lower frequencies) the resonant frequency becomes imaginary and linear in \(\Delta R\). The dominance of the imaginary part formally means that the width becomes large compared to the real part of the resonant frequency, which means that a resonance does not really exist. The crossover point is \(\Delta R = 0.01 \text{ nm}\) or \(\omega = 0.01 \text{ fs}^{-1}\) (\(\lambda = 180 \mu m\)) for the extended Drude model we employ. The dc conductivity \(\sigma\) employed in our model is about a factor of 3.5 too small. Thus our extended Drude model underestimates losses in the low frequency limit. However, the losses are slightly overestimated with respect to experiment at \(\lambda \sim 10 \mu m\) (see Fig. 2). Realistically, the
results in Fig. 6 should be accurate at least down to \( \omega = 0.1 \) fs\(^{-1} \) with the downturn for both curves becoming more severe at lower frequencies. Figure 7 shows the corresponding curves for the case \( k_z = 0.02166 \) nm\(^{-1} \), which is the \( m = 1 \) longitudinal resonance condition for \( L = 290 \) nm. Here also the \( \omega_{res} \sim \Delta R^{1/2} \) behavior occurs over the same range of \( \Delta R \) (and \( \omega \)) as in Fig. 6.

**IV. FDTD SIMULATION OF COAXIAL RING APERTURES**

The above CSP normal mode analysis largely explains the unusual dependence of the decreasing resonant frequency (or increasing resonant wavelength) with the narrowing of the coaxial rings, as well as predicts this dependence for rings much narrower than can be explored with numerical simulation. Numerical simulation, however, is necessary to supplement these predictions for several reasons: (i) The CSP dispersion analysis does not tell us how efficiently the CSP normal modes couple with the incident radiation, i.e., the intensity of these resonances. (ii) The resonant condition \( k_z L = m \pi \) is approximate. (iii) All resonant modes (SP, CSP, \( n \neq 1 \)) interfere naturally in a FDTD simulation, and should give more reliable predictions of transmission peaks, etc. (iv) The simulations include the full effect of the dielectric constants, including losses. In this section we confirm the theoretical predictions of the previous section and further assess the properties of the coaxial ring arrays through FDTD simulations.

The previous theoretical analysis assumes that the CSP enhancements are the result of the CSP normal modes of individual CR apertures. This would seem justified because the fields drop precipitously between apertures in the gold film (see Fig. 5), i.e., a tight-binding picture seems to apply. However, periodicity effects do exist between cylindrical apertures even for perfectly conducting (PC) films. To see how these peaks are affected by the periodicity of the lattice, in Fig. 8 we compare the transmission spectrum of 555 nm periodicity with that of 888 nm periodicity. The longest wavelength PSP for 888 nm periodicity should be at about 1350 nm. A peak does show up at 1333 nm, but it is hardly discernible in Fig. 8. (There is a larger peak at 1243 nm, but this is hardly discernible also.) The largest peak appears at 1004 nm, very close to the 1000 nm peak for 555 nm periodicity. (The intensity is down by a factor of 3, but this mainly reflects the reduction of the aperture density by a factor 2.56). Thus the position of this peak is hardly affected by periodicity. Baida et al.\(^{10–12} \) obtain the same result as long as the periodicity is greater than 300 nm (in silver). This strongly implies that the long wavelength enhancements depend only on the physics of the individual CR and the CSP modes uncoupled from the other CR’s.

Figure 8 also includes the spectrum for the Au film replaced by a perfectly conducting film with CR apertures of the same geometry. A peak appears at 808 nm (554 nm in the silica), well beyond the cutoff of 687 nm for the TE\(^1 \) mode. This peak is likely the result of the resonant coupling between the rings,\(^6 \) which occurs at a wavelength (in the medium) close to the lattice spacing, and occurs even for PC films. Our analysis does not preclude such effects also in the real metal, but the long wavelength enhancements and the singular behavior of the resonant wavelength as \( R_1 \) approaches \( R_2 \) are not related to such effects and explicitly

---

**FIG. 7.** The real and imaginary parts of the \( m = 1 \) \( k_z = 0.02166 \) nm\(^{-1} \) resonant frequency for the full dielectric constant of Eq. (1).

**FIG. 8.** The transmission spectra for the periodicities \( d = 555 \) nm and \( d = 888 \) nm for gold films, and for \( d = 555 \) nm for a perfectly conducting (PC) film. All other parameters are the same as our basic model.
depends on the negative dielectric constant of the real metal. To test our predictions on the dependence of the transmission peaks on \( R_1 \) and \( L \), we illustrate in Fig. 9 the transmission spectra for two different \( L \) values for \( R_1 = 50 \) nm and \( R_1 = 75 \) nm. The theoretical resonance positions, with the condition \( k_z L = m \pi \) assumed, are also indicated. The theory and simulations are in good agreement with respect to the movement and spacing of the peaks with \( R_1 \) and \( L \). The peaks in the simulation usually are slightly redshifted (by an average \( \sim 30 \) nm) from the theoretical positions, a feature that is also common for peaks associated with planar surface plasmon resonances. The peak positions at the longest wavelengths hardly change with \( L \) because they correspond to \( m = 0 \) and in this case \( k_z = 0 \) independent of \( L \). However, the second peaks (the third peak for \( R_1 = 50 \) nm, \( L = 290 \) nm, since the second peak here is the PSP peak), corresponding to \( m = 1 \), do significantly redshift for larger \( L \). This is consistent with the dispersion curves of Fig. 4 since here \( k_z = \pi / L \), and the resonant wavelength is a monotonically decreasing function of \( k_z \) (Fig. 4). Additionally, the PSP peak at 903 nm for our basic case hardly changes when \( L \) increases from 290 to 1000 nm, thus distinguishing PSP behavior from that of CSP’s. The intensity, however, is significantly depressed for \( L = 1000 \) nm, which is mainly the result of losses. Moreover, the results in Fig. 9 closely confirm the theoretical predictions of Sec. III.

We can also extract spatial fields and Poynting vectors (energy flow) from the simulations to examine the character of the solutions. The radial behavior of the fields in the CR we extract closely resemble those in Fig. 5. Figure 10 illustrates the real and imaginary parts of the Poynting vector (\( S \)) for our basic case at the PSP peak (903 nm) and at the CSP peak (1000 nm). The signature of a CSP resonance is a large imaginary part of the Poynting vector along the surfaces of the CR throughout the CR. This is observed in Fig. 10(b) at the CSP peak even in the middle of the CR, but not in Fig. 10(a) for the PSP peak, although there is a large \( \text{Im}(S) \) at the top and bottom surfaces. \( \text{Im}(S) \) is strongest at the inner metal-dielectric interface, although this is hard to see in the figure. Also, as a by-product, the real part of \( S \) is appreciable in the CR and decays only very slowly (due to losses) in both cases. The behavior of \( S \) is indicative of guided modes rather than evanescent modes in the CR, even at nonresonant frequencies. Thus the modes in the coaxial apertures influence the propagation through the CR even off-resonance and can lead to overall enhancements even for the PSP modes, by dramatically increasing the TE\(_1\) mode cutoff.

One of our most important results is that the resonant wavelength seems to be ever increasing as the CR is narrowed, behaving as \( \sim (R_2 - R_1)^{-1/2} \) at least in the optical and the near infrared. Of course in the \( R_1 = R_2 \) limit the transmission vanishes because the aperture vanishes, i.e., the exposed dielectric area vanishes. For a cylindrical aperture the transmission decays faster than the exposed area because the waves in the aperture are evanescent. For the CR apertures the modes are propagating, thus we would expect the transmission to fall off roughly as the exposed area. Table I gives scaled values of the \( m = 0 \) and \( m = 1 \) peak wavelengths scaled to \( (R_2 - R_1)^{-1/2} \) (i.e., \( \lambda_{sc} = \lambda_{peak}/(R_2 - R_1)^{-1/2} \)), and the transmission at the peak scaled to the exposed area \( [T_{sc} = \text{Transmission(FDTD)} \times \text{Area(FDTD cell)} / \text{Area(CR)}] \) for \( L = 290 \) nm. Results for \( R_1 \approx 85 \) nm are for FDTD simula-

![Image](https://example.com/image.png)
Table I. The m = 0 and m = 1 CSP transmission peak wavelengths, determined from the dispersion curves (Fig. 4), scaled to \((R_2-R_1)^{-1/2}\) [i.e., \(\lambda_{sc}=\lambda_{peak}/(R_2-R_1)^{-1/2}\)] and the transmitted intensity scaled to the exposed CR area \([T_{sc} = \text{Transmission(FDTD) \times Area(FDTD cell)/Exposed Area(CR)}]\) for various \(R_1\) values. Results for \(R_1 = 85\) nm are for FDTD simulations, while for \(R_1 > 85\) nm the results are from the dispersion analysis (Fig. 4).

<table>
<thead>
<tr>
<th>(R_1) (nm)</th>
<th>(L) (nm)</th>
<th>(\lambda_{sc}) (nm)(^{1/2})</th>
<th>(T_{sc})</th>
<th>(\lambda_{sc}) (nm)(^{1/2})</th>
<th>(T_{sc})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>290</td>
<td>6820</td>
<td>2.16</td>
<td>5790</td>
<td>0.99</td>
</tr>
<tr>
<td>50</td>
<td>290</td>
<td>7071</td>
<td>4.01</td>
<td>5353</td>
<td>2.14</td>
</tr>
<tr>
<td>75</td>
<td>290</td>
<td>7030</td>
<td>2.95</td>
<td>5550</td>
<td>4.70</td>
</tr>
<tr>
<td>85</td>
<td>290</td>
<td>6801</td>
<td>2.12</td>
<td>5170</td>
<td>4.02</td>
</tr>
<tr>
<td>90</td>
<td>290</td>
<td>6596</td>
<td>2.07</td>
<td>4718</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>290</td>
<td>6481</td>
<td>4.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>99.9</td>
<td>290</td>
<td>8985</td>
<td>5202</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The singular behavior of the cutoff wavelength holds as long as 
the length of the CR. The peak transmission intensities decrease no faster than the exposed CR area, enabling a super extraordinary transmission at very long wavelengths. The transmission intensities at the peaks, in fact, are typically 2–4 times the incident flux on the exposed portion of the CR.

We propose the following possible physical argument for the CSP enhancements and their dependence on ring geometry. CSP’s are “launched” analogously to PSP’s. PSP’s are launched when the photon can exchange planar momentum to the SP equal to a characteristic momentum of the two-dimensional (2D) lattice. Under periodic boundary conditions this sets up standing wave PSP’s. The analogy to the periodic boundary condition azimuthally is the requirement of single-valuedness, and this is characterized by the \(\phi\) dependence \(\sim \exp(\pm \sin \phi)\), which can be interpreted as the CSP exchanging \(n\) units of angular momentum with the cylinder. If a CSP also propagates in the \(z\) direction, standing wave boundary conditions normally require \(k_zL=m\pi\), where \(k_z\) is the characteristic unit of momentum which is associated with the length of the aperture and exchanged with the CSP. Together \(n\) and \(k_z\) describe the CSP motion on a cylindrical surface as \(k_{sp}\) describes the motion of a PSP on a flat surface. When \((n,k_z)\) has a characteristic value of the cylinder (integral \(n, k_zL=m\pi\)) a CSP is launched. When the ring is narrow, the inner and outer CSP’s not only have the same angular momentum, but they have nearly the same angular velocity, meaning that the two CSP’s have more coherence. A complementary view is that the wave functions of the two CSP’s overlap increasingly as the ring narrows, as seen in Fig. 5.

The structures we study are difficult to fabricate in the optical range because of the high aspect ratios involved, and also because silver and gold, which are the metals that give the best SP response (i.e., low losses), are very inert. They are difficult to process in such dimensions. Electron and ion beam lithography, liftoff, dry etch, electroplating, and ion beam milling are potential techniques that could be used to fabricate such structures. Recently there has been significant progress in fabricating these coaxial structures. As previously mentioned, Salvi et al.\(^{13}\) employed electron-beam lithography and gold liftoff to form 330 nm diameter coaxial aperture arrays in 140 nm gold films on glass. However, they did not measure the structures in the near infrared range where we would have expected the CSP peak. Very recently Orbons et al.\(^{32}\) using ion-beam lithography, fabricated similar arrays on 70–190 nm thick silver films on a glass substrate and measured the transmission up to 1700 nm, which includes the region where the \(m = 0\) CSP peak should occur. (We are presently analyzing these results.) To date both of these experiments\(^{13,32}\) have been limited to rather low aspect ratios for the apertures.

While the present study was restricted to the optical and IR regimes, beyond which losses would dominate, one could exploit the features in other regimes (THz, RF) by the development of metamaterials\(^{33}\) with negative (effective) dielectric constants and low losses containing such CR structures. The properties we found should have a dramatic effect on device design and hopefully motivate experiments to fabricate (i.e., through ion beam milling, dry etch, electroplating, or related techniques) and explore these structures, as a start in this direction has already occurred.\(^{13,32}\)
ACKNOWLEDGMENTS

This work was partially supported by the Office of Naval Research. Computations were carried out under the Department of Defense High Performance Computation Modernization Project.