

Theory of light transmission through an array of rectangular holes

A. Mary,¹ Sergio G. Rodrigo,² L. Martín-Moreno,² and F. J. García-Vidal¹

¹*Departamento de Física de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*

²*Departamento de Física de la Materia Condensada-ICMA, Universidad de Zaragoza, E-50009 Zaragoza, Spain*

(Received 14 June 2007; revised manuscript received 24 August 2007; published 13 November 2007)

In a two-dimensional array of rectangular holes perforated on a metallic film, two mechanisms leading to enhanced transmission of light operate: excitation of surface plasmon polaritons (SPPs) and localized resonances that are also present in single holes. In this paper, we analyze theoretically how the two mechanisms evolve and mix when the period of the array is varied. We also demonstrate that absorption in the metal is the main limiting factor for the SPP-based enhanced transmission.

DOI: [10.1103/PhysRevB.76.195414](https://doi.org/10.1103/PhysRevB.76.195414)

PACS number(s): 42.25.Bs, 73.20.Mf, 42.79.Dj, 78.66.Bz

I. INTRODUCTION

The fundamental study of the transmission properties of subwavelength apertures has received a lot of attention since the pioneering experiments by Ebbesen *et al.*¹ In this reference, they reported an extraordinary optical transmission (EOT) through a two-dimensional (2D) array of subwavelength circular holes perforated on a metallic film. This resonant transmission, appearing at wavelengths close to the period of the array, has been theoretically assigned to the excitation of surface plasmon polaritons (SPPs).² Several works have recently shown that the shape of the subwavelength aperture also has a great importance in the transmission spectra.³⁻⁷ These studies demonstrated that elliptical or rectangular holes influence dramatically the resonant wavelength, the polarization, and the final transmittance. On the other hand, very recent experimental and theoretical works have shown that even a single rectangular hole exhibits a localized transmission resonance emerging at around the cutoff wavelength λ_c of the hole waveguide.⁸⁻¹⁰

The aim of this paper is to analyze in detail the interplay between these two types of transmission resonances (SPP and cutoff resonances) that are operating in a 2D array of rectangular holes. Figure 1 shows schematically the system under study: an infinite array of rectangular apertures of sides a_x and a_y perforated on a freestanding silver film of thickness h . The structure is illuminated by a p -polarized plane wave at normal incidence (i.e., incoming electric field points along the short edge of the holes). In our calculations, we use the same geometrical parameters as those of the experiments carried out in Ref. 7: $a_x=200$ nm, $a_y=260$ nm, and $h=400$ nm. The period of the array, d , will be varied between 500 and 900 nm.

The paper is organized as follows. In Sec. II, the theoretical formalism based on the modal expansion of the electromagnetic (EM) fields is outlined. Section III presents the results obtained by this approximated method in comparison with quasixact numerical calculations done with a finite-difference time-domain (FDTD) code. By analyzing the dependence of the transmission properties with the period of the array, the physical origin of the different transmission resonances appearing in 2D hole arrays is revealed. EM-field maps associated with these transmission resonances are also depicted in this section. Additionally, a study of the effect of

absorption on the transmission spectra is carried out. Finally, our conclusions are presented in Sec. IV.

II. THEORETICAL FORMALISM

In this section, we briefly outline the theoretical formalism used for calculating the transmittance through the structure (a detailed account of this method can be found in Ref. 11). This framework is based on a modal expansion of the EM fields in each region. In reflection (I) and transmission (III) vacuum regions, electric fields are expressed in terms of plane waves $|\mathbf{k}\sigma\rangle$, characterized by the in-plane component of the wave vector \mathbf{k} and the polarization σ as

$$|\mathbf{E}_I(z)\rangle = |\mathbf{k}_0\sigma_0\rangle + \sum_{\mathbf{k}\sigma} r_{\mathbf{k}\sigma} |\mathbf{k}\sigma\rangle, \quad (1)$$

$$|\mathbf{E}_{III}(z)\rangle = \sum_{\mathbf{k}\sigma} t_{\mathbf{k}\sigma} |\mathbf{k}\sigma\rangle, \quad (2)$$

where $r_{\mathbf{k}}$ and $t_{\mathbf{k}}$ are the reflection and transmission coefficients. Inside the holes, EM fields are expanded in terms of the TE and TM waveguide eigenmodes $|\alpha\rangle$ as

$$|\mathbf{E}_{II}(z)\rangle = \sum_{\alpha} |\alpha\rangle [A_{\alpha} e^{iq_{z\alpha}z} + B_{\alpha} e^{-iq_{z\alpha}z}], \quad (3)$$

where $q_{z\alpha}$ represents the propagation constant of mode α and A_{α} and B_{α} are its corresponding expansion coefficients. The magnetic field is derived straightforwardly. Since the metal conductance depends on the frequency of incident light,

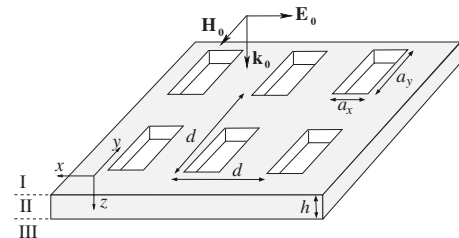


FIG. 1. Schematic picture of a square array of rectangular holes of sides a_x and a_y perforated on a freestanding silver film of thickness h . Parameter d defines the period of the array. The apertures are illuminated by a p -polarized plane wave at normal incidence.

surface-impedance boundary conditions (SIBCs) are imposed on the metallic boundaries, except on the vertical walls of the holes which are treated as perfect conductor surfaces. As we will see below, to consider only the fundamental TE eigenmode (TE₀₁) in the modal expansion within the holes gives accurate results for the transmittance spectrum. From now on, this mode is labeled as 0 mode. This approximation will allow us to write analytical expressions for the transmittance which highlight the underlying physics of the EOT phenomenon.

In the modal expansion of the EM fields inside the holes, it is convenient to relate the expansion coefficients with the quantities $E=A_0(1-Z_sY_0)+B_0(1+Z_sY_0)$ and $E'=A_0(1-Z_sY_0)e^{iq_z h}+B_0(1+Z_sY_0)e^{-iq_z h}$, which are the 0-modal amplitudes of the electric field at the input and output sides of the holes, respectively. $Z_s=1/\sqrt{\epsilon(\lambda)}$ is the surface impedance of the metal. In our study, the dielectric function $\epsilon(\lambda)$ of silver is taken from the tables of Ref. 12. $Y_{TE}=q_z/k_0$ corresponds to the admittance of the 0 mode with $k_0=2\pi/\lambda$. Realistic values for the propagation constant of the 0 mode, q_z , and for the effective side, a_y^{eff} , are incorporated into the formalism by using the effective index method.¹³ By analyzing the location in which q_z changes from being a real quantity to a purely imaginary magnitude, we can calculate λ_c within this approach. For the particular set of geometrical parameters of the holes forming the 2D array in Ref. 7, $\lambda_c=695$ nm.

By matching the EM fields at the interfaces, we end up with a set of two coupled linear equations for $\{E, E'\}$,¹¹

$$(G - \Sigma)E - G_V E' = I_0,$$

$$(G - \Sigma)E' - G_V E = 0. \quad (4)$$

The different terms of these equations have a simple physical interpretation and it is straightforward to find their expression. I_0 measures the overlap between the incident plane wave and the 0 mode inside the hole,

$$I_0 = \frac{4\sqrt{2}}{i\pi(1+Z_s)} \frac{a_x a_y}{d_x d_y}. \quad (5)$$

The term G_V in Eq. (4) controls the coupling between the input and output sides of the holes,

$$G_V = \frac{2iY_0 e^{iq_z h}}{e^{2iq_z h}(1+Z_sY_0)^2 - (1-Z_sY_0)^2}. \quad (6)$$

The expression for the *self-energy* Σ is given by

$$\Sigma = iY_0 \frac{e^{2iq_z h}(1+Z_sY_0) + (1-Z_sY_0)}{e^{2iq_z h}(1+Z_sY_0)^2 - (1-Z_sY_0)^2}. \quad (7)$$

Finally, the EM coupling between the holes forming the 2D array is mediated by the term G , which can be expressed as

$$G = \frac{ia_x a_y}{2d_x d_y} \sum_{l=-\infty}^{l=+\infty} \sum_{m=-\infty}^{m=+\infty} \frac{k_0(k_0 + Z_s k_z) - k_m^2}{(k_z + Z_s k_0)(k_0 + Z_s k_z)} \text{sinc}^2\left(\frac{k_l a_x}{2}\right) \times \left[\text{sinc}\left(\frac{k_m a_y + \pi}{2}\right) + \text{sinc}\left(\frac{k_m a_y - \pi}{2}\right) \right]^2, \quad (8)$$

with $k_l = \frac{2\pi}{d_x} l$, $k_m = \frac{2\pi}{d_y} m$, $k_p = \sqrt{k_l^2 + k_m^2}$, and $k_z = \sqrt{k_0^2 - k_p^2}$. For the rectangular hole geometry, the normalized-to-area transmittance T (transmission normalized to the amount of light impinging on the area occupied by the holes) is written as the following analytical expression:

$$T = \frac{a_y^{\text{eff}} k_0}{4a_y} \sum_{l=-\infty}^{l=+\infty} \sum_{m=-\infty}^{m=+\infty} \left[\frac{k_z}{|k_p|^2} \left(\frac{|k_l|^2}{|k_z + Z_s k_0|^2} + \frac{|k_m|^2}{|k_0 + Z_s k_z|^2} \right) \right] \times \left| \text{sinc}\left(\frac{k_l a_x}{2}\right) \left[\text{sinc}\left(\frac{k_m a_y + \pi}{2}\right) + \text{sinc}\left(\frac{k_m a_y - \pi}{2}\right) \right] E' \right|^2. \quad (9)$$

III. RESULTS AND DISCUSSION

A. Numerical results and origin of extraordinary optical transmission

Figure 2(b) shows the normalized-to-area transmittance spectra corresponding to the geometrical parameters of Ref. 7, calculated with the theoretical formalism explained above. Different curves correspond to different periods of the 2D square array (ranging from $d=500$ nm to $d=900$ nm). To compare them with results of a quasixact numerical framework, Fig. 2(a) depicts the corresponding transmission spectra obtained with the FDTD method.¹⁴ It is clear that our simplified formalism is able to capture accurately the main features observed in the FDTD spectra. The locations and linewidths of the several peaks are well reproduced within our approach. The heights of the transmission peaks are higher in our approach, mainly due to the fact that absorption within the vertical walls of the holes is neglected. It is also important to note that the theoretical results of panel (b) are in very good agreement with the experimental data.⁷

Now, we concentrate on analyzing the physical origin of the different transmission peaks appearing in Fig. 2(b) by just studying the simple set of Eq. (4). As in a typical resonant phenomenon, the transmittance is governed by the behavior of the denominator, which in Eq. (4) is $(G - \Sigma)^2 - G_V^2$. In Fig. 3, we show the dependence of both $|G - \Sigma|$ and $|G_V|$ versus wavelength for two limiting values of d , $d=500$ nm and $d=800$ nm [notice that G_V does not depend on d , see Eq. (6)]. Interestingly, the spectral locations of the transmission peaks in Fig. 2(b) are marked by the cuts between $|G - \Sigma|$ and $|G_V|$. It is worth comparing the behavior of $|G - \Sigma|$ for the periodic arrays with $d=500$ and 800 nm with the corresponding $|G - \Sigma|$ for a single rectangular hole (dot-dashed line in Fig. 3). As reported in Refs. 9 and 10, the transmittance through a single rectangular hole is also governed by a set of linear equations such as Eq. (4) but with a different G term [the sum over diffractive modes in Eq. (8) is replaced by an

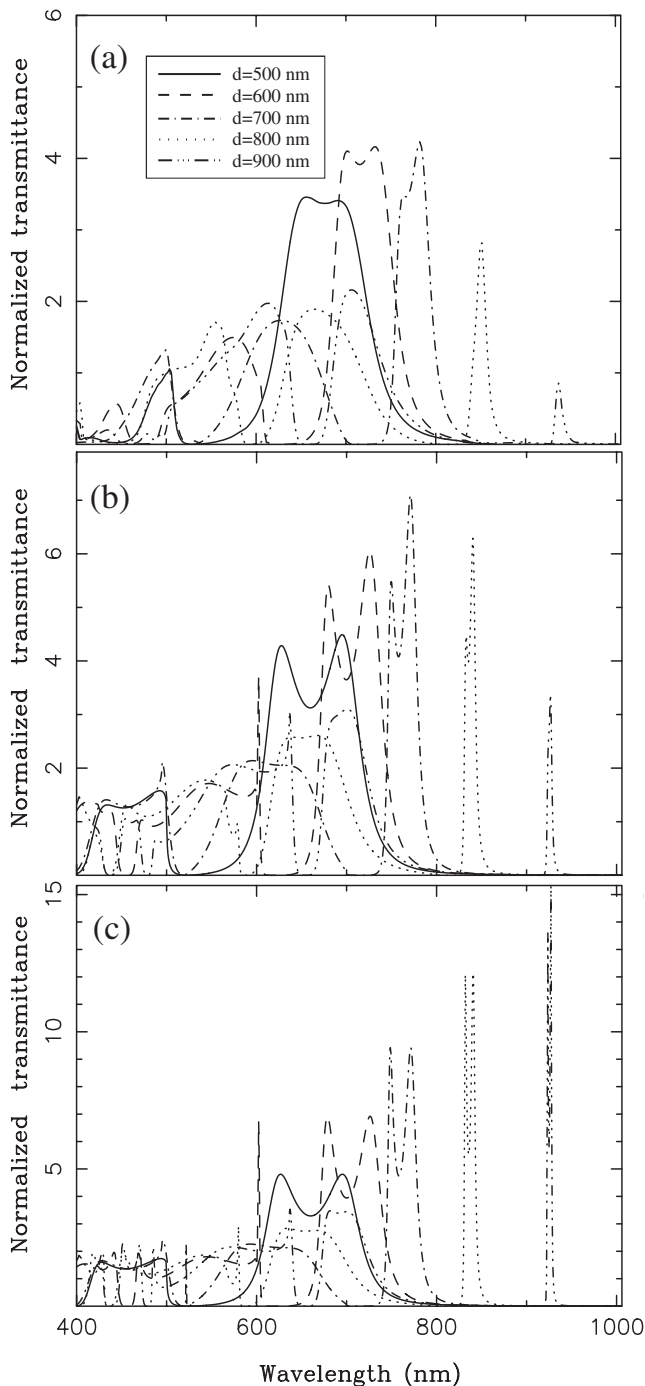


FIG. 2. (a) Normalized-to-area transmittances calculated with the FDTD method for different values of the lattice period d . Panels (b) and (c) show the normalized-to-area transmittances calculated with our modal expansion method. In panel (b), the dielectric function is that of silver, whereas in (c) absorption of silver is neglected [$\text{Im } \epsilon(\lambda) = 0$].

integral over the continuous spectrum of plane waves]. For $d = 500$ nm, $|G - \Sigma|$ (full line) at the wavelength region of interest (near $\lambda = 700$ nm) is close to the single hole counterpart. It is expected then that the nature of the transmission resonances will be similar for a 2D array and for an isolated rectangular hole. However, there is a difference between the

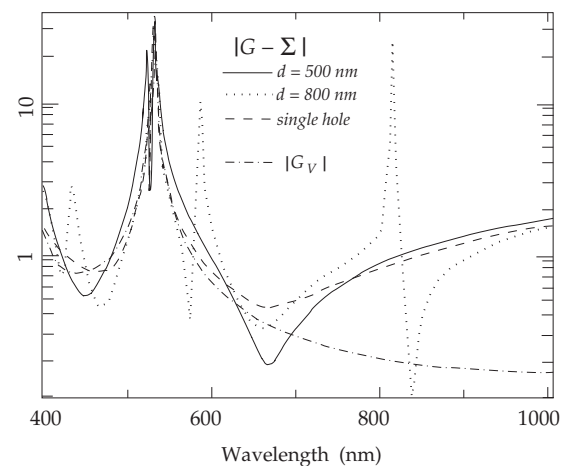


FIG. 3. Absolute values of $|G - \Sigma|$ and $|G_V|$ terms (in logarithmic scale) for two different lattice constants of the hole array ($d = 500$ nm and $d = 800$ nm) and also for a single hole.

single hole case and the 2D array for this value of d . Whereas in the 2D array, $|G - \Sigma| = |G_V|$ at two different wavelengths (leading to two transmission peaks); for the single hole, there is only one transmission peak appearing at a wavelength in which the difference between $|G - \Sigma|$ and $|G_V|$ is minimal. The reason of this distinct behavior stems from the Fabry-Pérot nature of this transmission resonance. In a Fabry-Pérot cavity, the spectral locations of the transmission resonances depend strongly on the reflectivity at the boundaries of the cavity. Our results indicate that the presence of a 2D array modifies the reflectivity of the metallic interface when compared to the single hole case, leading to the appearance of two transmission peaks. For $d = 800$ nm (dotted line), $|G - \Sigma|$ present additional features located at $\lambda \approx 600$ nm and $\lambda \approx 800$ nm. These correspond to zeros of the denominator of G [see Eq. (8)], appearing at the condition $k_z + Z_s k_0 = 0$. This condition is nothing else than the equation for the excitation of a SPP on a noncorrugated (no holes) surface of a metal film within the SIBC approach. Note that the cuts between $|G - \Sigma|$ and $|G_V|$ appear at wavelengths slightly larger than this condition. Therefore, the character of the two transmission peaks for $d = 800$ nm emerging at $\lambda \approx 830$ nm will be quite similar to a SPP. The two transmission peaks are associated with the symmetric and antisymmetric combinations of the two SPPs at the two surfaces which are evanescently coupled through the holes, as explained in Ref. 2.

As a general conclusion as regards the nature of the transmission resonances appearing in 2D arrays, we could state that the two mechanisms leading to EOT (localized resonance and SPP based) cannot be simply separated as done in previous studies.^{15,16} Even for the shortest and the longest period considered here, the transmission resonances benefit from both mechanisms. Therefore, these transmission resonances have a hybrid character. Hybrid resonances have already been observed and studied previously in lamellar^{17,18} and bottle-shaped gratings.¹⁹

The physical picture described above is reinforced when looking at the evolution as a function of d of the electric-field

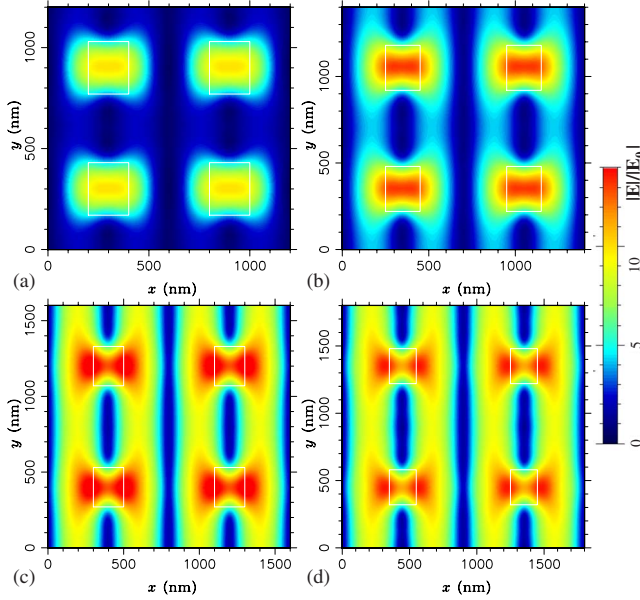


FIG. 4. (Color online) Electric-field amplitude calculated at the resonant wavelength and evaluated at the $z=0^-$ plane for (a) $d=600$ nm, (b) $d=700$ nm, (c) $d=800$ nm, and (d) $d=900$ nm. The white lines mark the positions of the holes. The E -field magnitude is normalized to the incident one.

patterns associated with the resonant process. In Fig. 4, we plot the E -field amplitude at resonance (normalized to that of the incident plane wave) evaluated at the $z=0^-$ interface. Four periods are considered here: $d=600$, 700, 800, and 900 nm. The electric-field amplitude for $d=600$ nm is mainly concentrated over the holes corresponding to the excitation of a localized resonance (cutoff resonance) inside the holes. As d is increased, the character of the resonance changes gradually. For $d=800$ and 900 nm, the E -field intensity maxima are along the ridges of the holes corresponding to a SPP wave propagating in the x direction. For $d=700$ nm, there is a mixing between the two mechanisms as fingerprints of the SPP wave begin to emerge at the edges of the holes. The near field distribution exhibits an intermediate character between the two limiting behaviors (SPP and cutoff resonances).

B. Effect of absorption

Finally, we would like to address the question of why there is a kind of optimum value for d when looking at the evolution of the transmission peaks with the period of the array [see Figs. 2(a) and 2(b)]. In our calculations, this optimum d is around 650–700 nm, close to the cutoff wavelength, 695 nm. Naively, this could imply that the optimum d appears when the resonant wavelengths of the two mechanisms (SPP and cutoff resonance) coincide. However, panel (c) of Fig. 2 demonstrates that the explanation is more complex. If the absorption in the metal is neglected, the heights of the transmission peaks grow with d like d^2 , as it would correspond to a perfect transmission (100%) per unit cell.²⁰ Therefore, EOT associated with SPP excitation is only lim-

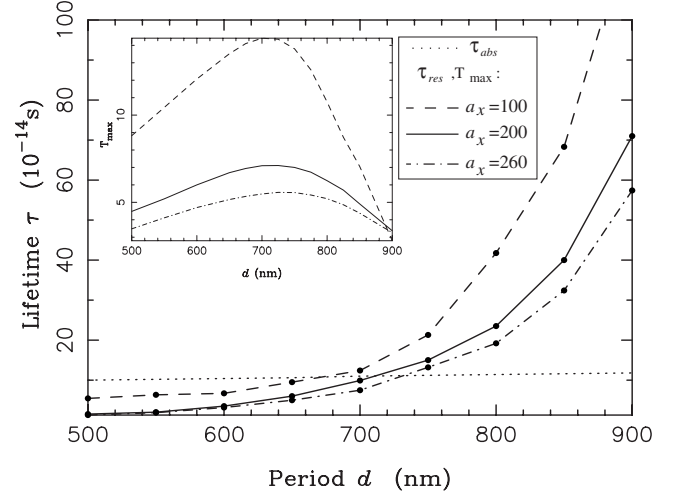


FIG. 5. Lifetimes of the resonant process τ_{res} and for absorption, τ_{abs} (dotted line) versus period of the hole array for $a_x=100$ nm (dashed line), $a_x=200$ nm (solid line), and $a_x=260$ nm (dot-dashed line). Inset: transmittance calculated at resonance versus d .

ited by absorption. As explained above, absorption along the walls of the holes is not taken into account in our approximated model. However, we have checked that for all d analyzed in this study, E -field intensity maxima are located at the horizontal metallic surfaces, where SIBCs are imposed within our modeling. Therefore, to consider only absorption on top and bottom surfaces of the metallic film is a reasonable approximation when analyzing the evolution of the heights of the transmission peaks as a function of d .

Within this approach, it is worth defining two different lifetimes operating during the transmission process. By looking at the linewidth of the transmission peaks with no absorption [Fig. 2(c)], we can extract the lifetime associated with the resonant process, τ_{res} . This quantity is depicted in Fig. 5 as a function of the period of the array for three values of a_x (200, 100, and 260 nm). As expected, when d is increased, resonant lifetime grows rapidly. On the other hand, absorption introduces another time into the problem. From the knowledge of $\epsilon(\lambda)$, we can estimate the time taken for a photon to get absorbed, τ_{abs} . This lifetime is almost independent of λ , as shown in Fig. 5 (dotted line). It is expected that optimum d would appear where $\tau_{res}(d) \approx \tau_{abs}(d)$. The line of reasoning leading to this naive rule is the following. When $\tau_{res}(d)$ is much smaller than $\tau_{abs}(d)$, photons are mainly transmitted and they are not absorbed by the metal. Absorption plays a minor role in the transmission process and the normalized-to-area transmittance at resonance increases when d is increased, as seen in Fig. 2(c). In the other limit [$\tau_{abs}(d)$ being much smaller than $\tau_{res}(d)$], photons are absorbed by the system before the resonance is built up. As $\tau_{res}(d)$ grows rapidly with d , a decrease of the transmittance at resonance versus d is expected to occur in this limit. When interpolating between these two limits, it is clear that the curve displaying the transmittance at resonance versus d should present a maximum for an optimum value of d . As $\tau_{res}(d)$ evolves very rapidly with d whereas $\tau_{abs}(d)$ is almost independent of d , optimum d should appear close to the con-

dition $\tau_{abs}(d)=\tau_{res}(d)$. Figure 5 demonstrates that this last condition marks the location of the optimum d for the three different values of a_x analyzed. Therefore, we can safely conclude that the physical origin of the optimum d observed in our calculations stems from the absorption present in the metallic film.

Note that in the experiments,⁷ the finite size of the hole array introduces a third lifetime associated with the spatial extension of the array, τ_{size} .¹¹ As an absorption in the metal, it also acts as a limiting factor in the final transmittance. If the number of holes is large enough, τ_{size} is greater than τ_{abs} and then the limiting factor would be the absorption by the metal, as in the case of an infinite array. However, if the array is very small, τ_{size} would be smaller than τ_{abs} and finite size effects would control the optimum d . Then, the cut between $\tau_{res}(d)$ and τ_{size} would mark the location of optimum d . If this is the case ($\tau_{size} < \tau_{abs}$), then the cut would appear at a shorter d . This seems to be the case in the experiments reported in Ref. 7 as the experimental optimum d is of the order of 600 nm, instead of a value close to 700 nm obtained from our calculations for an infinite array of holes (see Fig. 5).

IV. CONCLUSIONS

In conclusion, we have analyzed theoretically the interplay between two different mechanisms that enhance the transmission of light through a 2D array of rectangular holes: SPP-based extraordinary transmission and enhanced transmission assisted by the excitation of a localized resonance, spectrally located at the cutoff wavelength of the hole waveguide, λ_c . We have shown that when $d < \lambda_c$, the transmission resonance mainly has a localized nature whereas for $d > \lambda_c$, SPP governs the transmittance through the structure. We have also demonstrated that in this last case, resonant transmission is mainly limited by the absorption in the metal.

ACKNOWLEDGMENTS

Financial support by the EU under Contract No. IST-FP6-034506 and Spanish MEC under Contract No. MAT2005-06608-C02 is gratefully acknowledged.

-
- ¹T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, *Nature (London)* **391**, 667 (1998).
- ²L. Martín-Moreno, F. J. García-Vidal, H. J. Lezec, K. M. Pellerin, T. Thio, J. B. Pendry, and T. W. Ebbesen, *Phys. Rev. Lett.* **86**, 1114 (2001).
- ³K. J. Klein Koerkamp, S. Enoch, F. B. Segerink, N. F. van Hulst, and L. Kuipers, *Phys. Rev. Lett.* **92**, 183901 (2004).
- ⁴R. Gordon, A. G. Brolo, A. McKinnon, A. Rajora, B. Leathem, and K. L. Kavanagh, *Phys. Rev. Lett.* **92**, 037401 (2004).
- ⁵H. Cao and A. Nahata, *Opt. Express* **12**, 3664 (2004).
- ⁶K. L. van der Molen, K. J. Klein Koerkamp, S. Enoch, F. B. Segerink, N. F. van Hulst, and L. Kuipers, *Phys. Rev. B* **72**, 045421 (2005).
- ⁷A. Degiron and T. W. Ebbesen, *J. Opt. A, Pure Appl. Opt.* **7**, S90 (2005).
- ⁸A. Degiron, H. J. Lezec, N. Yamamoto, and T. W. Ebbesen, *Opt. Commun.* **239**, 61 (2004).
- ⁹F. J. García-Vidal, Esteban Moreno, J. A. Porto, and L. Martín-Moreno, *Phys. Rev. Lett.* **95**, 103901 (2005).
- ¹⁰F. J. García-Vidal, L. Martín-Moreno, Esteban Moreno, L. K. S. Kumar, and R. Gordon, *Phys. Rev. B* **74**, 153411 (2006).
- ¹¹J. Bravo-Abad, F. J. García-Vidal, and L. Martín-Moreno, *Phys. Rev. Lett.* **93**, 227401 (2004).
- ¹²*Handbook of Optical Constants of Solids*, edited by E. D. Palik (Academic, Orlando, 1985).
- ¹³R. Gordon and A. G. Brolo, *Opt. Express* **13**, 1933 (2005).
- ¹⁴A. Taflove and S. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Artech House, Boston, 2000).
- ¹⁵Zhichao Ruan and Min Qiu, *Phys. Rev. Lett.* **96**, 233901 (2006); see comment by Cheng-ping Huang and Yong-yuan Zhu, arXiv:0706.0250v1 (unpublished).
- ¹⁶Wulin Jia and Xiaohan Liu, *Eur. Phys. J. B* **46**, 343 (2005).
- ¹⁷T. López-Rios, D. Mendoza, F. J. García-Vidal, J. Sánchez-Dehesa, and B. Pannetier, *Phys. Rev. Lett.* **81**, 665 (1998).
- ¹⁸D. Crouse, M. Arend, J. Zou, and P. Keshavareddy, *Opt. Express* **14**, 2047 (2006).
- ¹⁹Diana C. Skigin and Ricardo A. Depine, *Phys. Rev. E* **63**, 046608 (2001).
- ²⁰F. J. García de Abajo, R. Gómez-Medina, and J. J. Sáenz, *Phys. Rev. E* **72**, 016608 (2005).