Influence of the electron spill-out and nonlocality on gap plasmons in the limit of vanishing gaps

M Khalid, O Morandi, Emilien Mallet, P Hervieux, G Manfredi, Antoine Moreau, C Ciraci

Influence of the electron spill-out and nonlocality on gap plasmons in the limit of vanishing gaps

M. Khalid,[1,*] O. Morandi [◉],[2] E. Mallet,[3,4] P. A. Hervieux,[5] G. Manfredi [◉],[5] A. Moreau [◉],[3,4] and C. Ciracì [◉][1,†]

[1]*Istituto Italiano di Tecnologia, Center for Biomolecular Nanotechnologies, Via Barsanti 14, 73010 Arnesano, Italy*
[2]*Dipartimento di Matematica e Informatica U. Dini, University of Florence, via S. Marta 3, 50149, Florence, Italy*
[3]*Clermont Université, Institut Pascal (IP), BP 10448, F-63000 Clermont-Ferrand, France*
[4]*CNRS, UMR 6602, IP, F-63171 Aubière, France*
[5]*Université de Strasbourg, CNRS, Institut de Physique et Chimie des Matériaux de Strasbourg, UMR 7504, F-67000 Strasbourg, France*

We study the effect of electron spill-out and of nonlocality on the propagation of light inside a gap between two semi-infinite metallic regions. We compare the predictions of a local response model taking into account only the spill-out, to the predictions of a quantum hydrodynamic model able to take both phenomena into account. We show that only the latter is able to correctly retrieve the correct limit when the gap closes, while the local model suffers from undesirable features (divergence of the fields, overestimation of the losses). Finally, we show that, to a certain extent, the correct results can be retrieved using a simple local approach without spill-out or a conventional Thomas-Fermi approximation, but considering an effective gap width.

I. INTRODUCTION

The possibility to design extremely miniaturized devices, even optical circuits, has fostered research in plasmonics for decades. The most recent advances in fabrication processes have led to metallic nanostructures where extremely narrow gaps play a central role [1]. Metal-insulator-metal (MIM) structures support a mode with a particularly large effective index called the *gap plasmon* [2,3]. The insertion of MIM sections in finite systems can thus lead to resonances in exceptionally small volumes [4,5], while the quality factor of the resonance [6] and its absorption cross-section [7] are essentially preserved. The insulator gaps can be so small (typically smaller than 1 nm [8]) that many questions arise concerning the accuracy of the electromagnetic description at such scales. Indeed, at these scales the classical local response approximation (LRA), based on the Drude model, reaches its limits [9], as the spatial dispersion due to the repulsion between electrons inside the metal has a direct impact on the optical response of the system. Electron-electron interactions can be taken into account in a simple way in the framework of the hydrodynamic theory (HT) with hard-wall boundary conditions (i.e., no electron spill-out is allowed) by adding the Thomas-Fermi (TF) electron pressure term [10,11].

When dealing with the near-atomic subnanometer length scales, it becomes inevitable to consider the quantum nature of electrons, and the quantum-mechanical effects, such as electron spill-out and quantum tunneling, become important. The TFHT cannot address these phenomena properly, therefore more sophisticated theories are required for an accurate description of electron dynamics [12–14]. *Ab initio*

quantum-mechanical approaches, such as time-dependent density functional theory (TD-DFT), provide an appealing solution to such problems, although they are often computationally demanding, especially for large plasmonic structures [15–17].

Within the LRA framework, it is still possible to account for the tunneling by interposing between the metal regions a fictitious material described by an effective conductivity [18] or by using a finite-current boundary condition at the metal-insulator interface [19–21]. However, when the LRA is used to account for spill-out, nonphysical features (including a divergence of the associated electric field) arise [22,23]. Due to this artificial divergence of the field, the imaginary part of the propagation constant of the gap plasmon can be largely overestimated, which consequently results in very high mode-propagation losses.

In the past few years, a hybrid approach called quantum hydrodynamic theory (QHT)—which adds a ∇n-dependent correction (n being the charge density) to the TF kinetic energy—has been developed to study nonlocal and quantum effects (electron spill-out/tunneling) in plasmonic systems more accurately [24–28]. The QHT method intrinsically accounts for the electron spill-out and tunneling, but also spatial dispersion (nonlocality), which is relevant in the unidirectional plasmonic waveguides [29] and nonlinear optical dynamics [30,31]. This approach is based on an orbital-free description of quantum electronic systems, and its accuracy relies on the exactness of a noninteracting kinetic energy functional. Ideally, if the exact expression of such a functional were known, QHT would be as accurate as TD-DFT [32].

Here we apply the QHT approach to the gap plasmon in the vanishing gap limit, and we compare to the LRA model in the presence of spill-out. We show that the unphysical divergences predicted by the LRA lead to a large and artificial overestimation of the absorption when the gap plasmon

---

*muhammad.khalid@iit.it
†cristian.ciraci@iit.it

propagates, but that including spatial dispersion in the description of the electron gas makes such features disappear. In fact, the gap-mode propagation constant is influenced by the presence of longitudinal modes (due to nonlocality) inside the metallic region, which are determined by the hydrodynamic equations and, in particular, by the quantum pressure term. The correct limit, as the gap closes, is thus retrieved only when both the spill-out and nonlocality are taken into account.

In Sec. II, we detail the theoretical background of our study, first by deriving simplified equations corresponding to the LRA in the presence of spill-out, and then by detailing the differences with the QHT framework. In Sec. III, we discuss our numerical results by comparing the behavior of the effective index for the different approaches of interest (LRA, TFHT, LRA with spill-out, and QHT). Finally, in Sec. IV we discuss the implications of our work regarding the definition of the actual optical edges of a metal. We show that it makes sense to define an effective gap for the gap plasmon, allowing us to understand why in so many works the LRA without spill-out has been so successfully applied.

II. THEORETICAL FRAMEWORK

The propagation of the electric and magnetic fields, **E** and **H**, of the gap plasmon is described by Maxwell's equations:

$$\nabla \times \mathbf{E} = -\partial_t \mu_0 \mathbf{H},$$
$$\nabla \times \mathbf{H} = \varepsilon_0 \varepsilon_\infty \partial_t \mathbf{E} + \mathbf{J}, \quad (1)$$

where μ_0 and ε_0 are the vacuum permeability and permittivity, respectively; ε_∞ is the relative permittivity accounting for any dielectric local response. **J** is the current inside the electron gas. This current is given by $\mathbf{J} = -en\mathbf{u}$, where **u** is the velocity of the electron gas, considered as a fluid, e is the electron charge in absolute value, and n is the electron density. In this framework, the velocity satisfies the following equations [33,34]:

$$\partial_t n + \nabla \cdot (n\mathbf{u}) = \mathbf{0},$$
$$\partial_t \mathbf{u} + \gamma \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u}$$
$$= -\frac{e}{m}(\mathbf{E} + \mathbf{u} \times \mu_0 \mathbf{H}) - \frac{1}{m}\nabla \frac{\delta G[n]}{\delta n}, \quad (2)$$

where m is the electron mass, and the energy functional $G[n]$ contains all the internal energy of the electronic system (here the operator $\frac{\delta}{\delta n}$ indicates the functional derivative with respect to the density n). The last term in the second equation gives rise to a nonlocal contribution to the current **J**. Using the previous equations, we obtain the following evolution equation for the current:

$$\partial_t \mathbf{J} + \gamma \mathbf{J} - \frac{1}{e}\left(\frac{\mathbf{J}}{n}\nabla \cdot \mathbf{J} + \mathbf{J} \cdot \nabla \frac{\mathbf{J}}{n}\right)$$
$$= \frac{ne^2}{m}\mathbf{E} - \frac{\mu_0 e}{m}\mathbf{J} \times \mathbf{H} + \frac{en}{m}\nabla \frac{\delta G[n]}{\delta n}. \quad (3)$$

Neglecting higher-order terms and taking $n(\mathbf{r}, t) = n_0(\mathbf{r}) + n_1(\mathbf{r}, t)$ and $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0(\mathbf{r}) + \mathbf{E}_1(\mathbf{r}, t)$, with n_0 being the equilibrium charge density and $\mathbf{E}_0$ the electrostatic field, we

first obtain an equation for the dynamic part,

$$(\partial_t + \gamma)\mathbf{J} = \frac{e^2}{m}(n_0 \mathbf{E}_1 + n_1 \mathbf{E}_0)$$
$$+ \frac{e}{m}\left[n_0 \nabla\left(\frac{\delta G[n]}{\delta n}\right)_1 + n_1 \nabla\left(\frac{\delta G[n]}{\delta n}\right)_0\right], \quad (4)$$

where $(\frac{\delta G[n]}{\delta n})_0$ and $(\frac{\delta G[n]}{\delta n})_1$ are the zeroth- and first-order terms of the potential, respectively. Second, we get an equation linking the static electron density to the static electric field,

$$\nabla\left(\frac{\delta G[n]}{\delta n}\right)_0 + e\mathbf{E}_0 = 0, \quad (5)$$

where $\mathbf{E}_0$ and n_0 can be found using Gauss's law:

$$\nabla \cdot (\varepsilon_\infty \mathbf{E}_0) = -\frac{e}{\varepsilon_0}(n_0 - n^+), \quad (6)$$

n^+ being the positive background density.

Finally, combining Eqs. (1) and assuming a harmonic field oscillating at ω, we obtain

$$\nabla \times \nabla \times \mathbf{E}_1 - \varepsilon_\infty k_0^2 \mathbf{E}_1 = i\omega \mu_0 \mathbf{J}, \quad (7)$$

$$(\gamma - i\omega)\mathbf{J} = \frac{e^2}{m}(n_0 \mathbf{E}_1 + n_1 \mathbf{E}_0)$$
$$+ \frac{e}{m}\left[n_0 \nabla\left(\frac{\delta G[n]}{\delta n}\right)_1 + n_1 \nabla\left(\frac{\delta G[n]}{\delta n}\right)_0\right], \quad (8)$$

where $k_0 = \omega/c$.

Combining Eqs. (5) and (6), we obtain

$$\nabla\left[\varepsilon_\infty \cdot \nabla\left(\frac{\delta G}{\delta n}\right)_0\right] - \frac{e^2}{\varepsilon_0}(n_0 - n^+) = 0. \quad (9)$$

Note also that using Eq. (5), we can simplify Eq. (8) as

$$(\gamma - i\omega)\mathbf{J} = \frac{e^2}{m}n_0 \mathbf{E}_1 + \frac{en_0}{m}\nabla\left(\frac{\delta G[n]}{\delta n}\right)_1. \quad (10)$$

A. LRA in the presence of spill-out

We first derive the equations corresponding to the LRA [18,22,23]. They have the merit to allow a better grasp of the physics at play. The main assumption here is to neglect the pressure term, i.e., $G[n] \simeq 0$, making the description purely local.

The impact of nonlocality alone in the case of vanishing gaps (and hard-wall boundaries) has already been investigated in Ref. [35], where it has been shown that the effective index of the gap plasmon increases although staying finite (contrary to the diverging behavior of the LRA case without spill-out [36,37]). Yet, it does not tend to the expected value of the bulk metal refractive index either.

Starting from Eq. (10), we get

$$\mathbf{J} = \frac{1}{\gamma - i\omega}\frac{e^2}{m}n_0 \mathbf{E}_1. \quad (11)$$

The current can always be incorporated into Eqs. (1) as an effective polarization P_f, by simply writing $\partial_t P_f = J$. This finally allows us to introduce a local susceptibility by writing,

just as for a Drude model, that

$$\mathbf{P}_f = -\varepsilon_0 \frac{\frac{e^2 n_0(\mathbf{r})}{\varepsilon_0 m}}{\omega^2 - i\gamma\omega} \mathbf{E}_1 = \varepsilon_0 \chi_f \mathbf{E}_1, \qquad (12)$$

except that the susceptibility is proportional to the electron density and thus depends on the position here.

Now if we consider the propagation of a guided mode in the z direction, and assume that all fields are invariant in the x direction, Maxwell's equations in the p-polarization reduce to

$$\partial_y E_z - \partial_z E_y = i\omega\mu_0 H_x, \qquad (13)$$

$$\partial_z H_x = -i\omega\varepsilon_0 \varepsilon(y) E_y, \qquad (14)$$

$$-\partial_y H_x = -i\omega\varepsilon_0 \varepsilon(y) E_z, \qquad (15)$$

where $\varepsilon(y) = \varepsilon_\infty + \chi_f(y)$ is the local permittivity.

Combining the previous equations to get rid of H_x, we obtain

$$\partial_y \partial_z E_z - \partial_z^2 E_y = i\omega\mu_0 \partial_z H_x = k_0^2 \varepsilon(y) E_y, \qquad (16)$$

$$\partial_y^2 E_z - \partial_y \partial_z E_y = i\omega\mu_0 \partial_y H_x = -k_0^2 \varepsilon(y) E_z. \qquad (17)$$

Since we are looking for a guided mode, we take $\partial_z = ik_z$ so that we obtain

$$\left[k_z^2 - \varepsilon(y) k_0^2 \right] E_y = -ik_z \partial_y E_z, \qquad (18)$$

and finally we manage to keep only the E_z field to write the wave equation under the form

$$\partial_y^2 E_z + k_0^2 \, \varepsilon(y) \, E_z + \partial_y \left(\frac{k_z^2}{\varepsilon(y) \, k_0^2 - k_z^2} \partial_y E_z \right) = 0. \qquad (19)$$

We underline that such an approach is not self-consistent, as it requires us to choose arbitrarily the electron density. Putting aside the choice of electron density, our local model with spill-out is identical to the model used in [22] and yields perfectly similar results. In Ref. [22], the authors compute n_0 using a quantum-mechanical approach, whereas we use a simple model density to approximate the electron spill-out (details are given in the next section).

We have then converted this equation into a numerical problem using a finite-difference scheme in order to look for values of k_z in the complex plan, which allows us to make the determinant of the finite-difference matrix vanish. As such a problem is extremely unstable numerically, obviously because of the different scales involved (the gap being much smaller than the skin depth), we have used the exponentially decaying analytic solution inside the metal. This solution is known when $\varepsilon(y)$ is uniform and corresponds to the classical solution of Maxwell's equations considering Drude's model for the metal. Connecting the finite-difference problem with the analytic solution allows us to reduce the size of the problem and to improve the stability, finally allowing us to compute the wave vector of the gap plasmon, even for vanishing gaps. We compared it with a finite-element-based approach, and both are in almost perfect agreement (see the Appendix).

B. Self-consistent quantum hydrodynamic theory

Equation (9) with Eqs. (7) and (10) constitute the basis of the QHT [24,26] linear response and allow us to calculate the detailed electron gas dynamic at the metal surface. In this work, we employ the following approximation:

$$G[n] = T_{\mathrm{TF}}[n] + \lambda_{\mathrm{vW}} T_{\mathrm{vW}}[n, \nabla n] + E_{\mathrm{XC}}[n], \qquad (20)$$

where T_{TF} is the TF kinetic energy, and T_{vW} is the gradient-dependent correction, namely the von Weizsäcker term, to the TF functional, which allows us to take into account effects depending on the gradient of the electron density n and thus should not be neglected when spatially dependent densities are considered. E_{XC} is the exchange-correlation energy functional. The parameter λ_{vW} in front of the von Weizsäcker term can be related to the decay of the charge density from a metal surface, and hence controls the amount of the electron spill-out and tunneling. It has been shown in Ref. [26] that for QHT, λ_{vW} is inversely proportional to the decay of the charge density and can be related to the "equilibrium" and "induced" electron spill-out. In general, λ_{vW} is a free parameter of the kinetic functional that can be tuned depending on what the figure of merit is (i.e., plasmon energy, equilibrium density, asymptotic decay, tunneling, etc.). Most often, in the literature a value in the [1/9, 1] range is used for λ_{vW}, in which $\lambda_{\mathrm{vW}} = 1/9$ corresponds to a lower electron spill-out and a faster decay of n_0 from a metal surface, whereas $\lambda_{\mathrm{vW}} = 1$ gives a higher spill-out and a slower decay of the charge density. In general, using $\lambda_{\mathrm{vW}} = 1/9$ approximates well the Lindhard function for small k-vectors, while $\lambda_{\mathrm{vW}} = 1$ gives a good estimation for large k-vectors [38]. In fact, the von Weizsäcker correction in Eq. (20) is the first-order term in the expansion of the kinetic energy [39], and to construct a more generally valid functional one would need to consider higher-order terms (i.e., Laplacian dependence), which would introduce more free parameters [40].

The explicit expressions of the energy functional $G[n]$ and its functional derivative, i.e., $(\frac{\delta G}{\delta n})_1$, can be found in Ref. [26]. We further note that Landau damping is another important factor to consider, and it can also be incorporated into Eq. (10), as was done in Ref. [27] in the context of QHT. We expect that the Landau damping will result in increased damping inversely proportional to the gap size. Since the main objective of the present work is to analyze the impact of nonlocality and electron spill-out on the gap plasmons, we neglect Landau damping. It is informative to note that considering a spatially constant $n_0 = n_b$, n_b being the charge density of the bulk metal, and $G[n] = 0$ in Eq. (20) leads to the usual Drude-type relation (LRA) for the polarization, and using $T_{\mathrm{vW}} = E_{\mathrm{xc}} = 0$ yields $G[n] = T_{\mathrm{TF}}$, the standard hydrodynamic theory in the Thomas-Fermi approximation (TFHT).

In contrast to the conventional theories in plasmonics, QHT can efficiently take into account the nonlocal, electron spill-out, and tunneling effects, and it can provide the details of the microscopic distribution of the fields, which usually cannot be described by conventional approaches. The QHT is highly dependent on the spatially varying equilibrium charge density n_0, and consequently the approximation of the optical response of a nanoplasmonic system relies strongly on the description of n_0. Over the past few years, QHT has been

applied to probe linear plasmonic properties of a variety of individual nanoparticles as well as nearly touching nanostructures (i.e., in the tunneling regime) [26,41–43], and a fairly good agreement between the QHT and TD-DFT calculations has been reported [41]. Very recently, electron spill-out effects in singular metasurfaces have been analyzed using QHT, showing that the spill-out effectively blunts the sharp singularities [44]. QHT has also been applied to study nonlinear optical properties of thick metal films, and it has been shown that there exist spill-out-induced resonances that can be exploited to achieve very large second-harmonic-generation efficiencies [31]. These resonances cannot be excited in the LRA and TFHT, as electron spill-out is overlooked in these approximations. Therefore, it becomes important to take into account the nonlocal and quantum effects when dealing with plasmonic structure with subnanometer geometrical features. Despite the fact that QHT gives a very good prediction of the plasmon resonances by adding a ∇n-dependent correction to the kinetic energy, it also suffers some limitations. For example, it predicts unexpected modes between the surface and bulk plasmon frequencies lying in the far tail region of the exponentially decaying charge density. Very recently, a development in the QHT theory based on the Laplacian-level kinetic energy functionals has been presented [40]. The introduction of Laplacian-dependent charge density results in more robust numerical solutions, but its implementation becomes more complex. Since in the present study we only consider very small gaps, the tail of the charge density never extends to the problematic region. Therefore, we expect that the QHT approach in the limit of the von Weizsäcker approximation used herein gives us results with the same level of accuracy and with the advantage of avoiding additional numerical complexity, arising from the Laplacian-level functionals. We adopt the jellium approximation [45] in implementation of the QHT, which assumes that the electrons in a metal are confined by a constant positive background charge $n^+ = (\frac{4}{3}\pi r_s^3)^{-1}$, where r_s indicates the Wigner-Seitz radius ($r_s = 4$ for Na and $r_s = 3$ for Ag).

To compute the gap-plasmon mode and its dispersion relation, we first compute the space-dependent equilibrium charge density using Eq. (9) and then solve Eqs. (7) and (10) assuming a solution of the type $\mathbf{E}(\mathbf{r}) = \tilde{\mathbf{E}}(y) e^{ik_{gp}z}$ (with similar expressions for $\mathbf{J}$), where k_{gp} is the propagation constant of the gap-plasmon mode. For a given angular frequency ω, we solve this system of equations using the finite-element method to compute the associated propagation constant $k_{gp}(\omega)$, as well as the mode $\tilde{\mathbf{E}}(y)$.

III. RESULTS

Let us consider a MIM configuration as schematically shown in Fig. 1, that is, a dielectric gap g sandwiched between two semi-infinite metallic regions. For simplicity, we use air as the dielectric. We examine the impact of nonlocality and the electron spill-out on the gap plasmon, the unique mode guided in p-polarization in the structure and propagating along the z direction. More precisely, we compute the effective index $n_{\text{eff}} = \frac{k_z}{k_0}$ of the gap plasmon according to the different models described above, and we compare the results, in particular in the gap size range (<1 nm) where the charge densities from
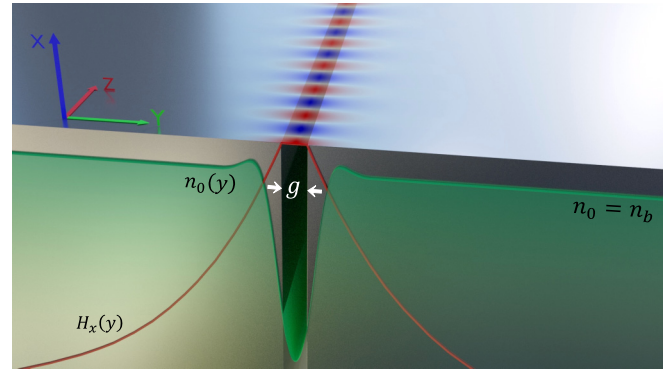

FIG. 1. Visual representation of overlapping of the equilibrium charge densities $n_0(y)$ due to electron spill-out in the dielectric gap g between two semi-infinite bulk metals. Deep inside the bulk metal, n_0 is equal to the constant density of the bulk metal n_b. The magnetic mode profile $H_x(y)$ of a p-polarized light propagating along the z-direction is also shown. The structure is invariant along the x- and z-directions.

the two surfaces significantly overlap, as schematically shown in Fig. 1. We begin with the LRA with spill-out for simple Drude metals, i.e., sodium (Na), which we compare with the LRA and the TFHT approximation without spill-out as well as with the QHT. Finally, we consider the case of silver (Ag) in the QHT framework, and we compare it to the commonly used LRA and TFHT without spill-out.

A. LRA with spill-out

We begin with the simplified case presented above in which the equilibrium electron density $n_0(\mathbf{r})$ can be described by an exponentially decaying profile with the following analytical function:

$$n_0(y) = n_b \left[\frac{1}{1 + e^{\kappa(y + \frac{g}{2})}} + \frac{1}{1 + e^{-\kappa(y - \frac{g}{2})}} \right], \quad (21)$$

where the parameter κ defines the spill-out or the decay of the charge density from the metal surface. In what follows, we set $\kappa \simeq (0.7a_0)^{-1}$, with a_0 being the Bohr radius, which we get by fitting the analytical function with the decay of the self-consistent charge density [see Eq. (9)], considering $\lambda_{\text{vW}} = 1/9$. We take $r_s = 4$, $\varepsilon_\infty = 1$, and $\hbar\gamma = 0.16$ eV, values that correspond to Na. Although it is impossible to work with Na in practical applications, due to its high reactivity, it presents an ideal Drude response as the impact of interband transitions can be omitted, thus offering a convenient theoretical study platform.

In Fig. 2, we compare the results of our simplified LRA with spill-out against the commonly used LRA without any spill-out. The input equilibrium charge-density profiles used in both approaches are plotted in Fig. 2(a) for $g = 1$ nm. In Figs. 2(b) and 2(c), we compare the behavior of the real and imaginary parts of the effective mode index. As expected, for large gaps ($g > 1$ nm) spill-out can be neglected; the LRAs with and without electron spill-out agree very well. The spill-out has a noticeable impact on the real part of the effective index for gaps that are typically around 1 nm or less: it increases the effective index, signaling a plasmonic slow
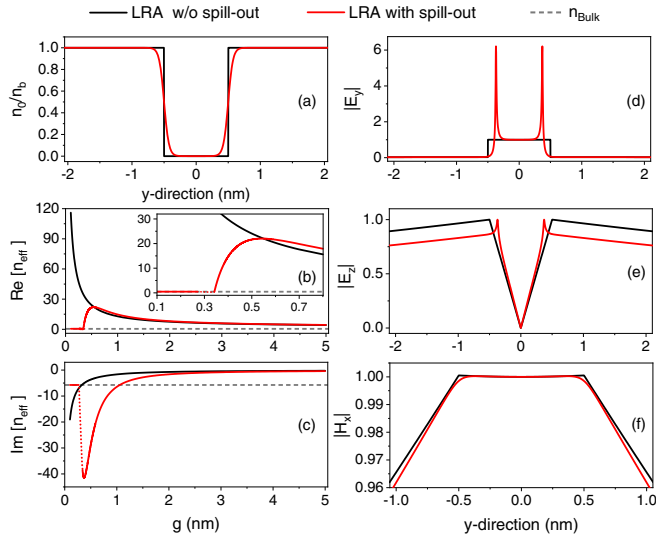
FIG. 2. A comparison between the local response approximation (LRA) with and without electron spill-out from the metal surfaces. (a) Equilibrium electron density n_0 normalized with the bulk density n_b for $g = 1$ nm. The (b) real and (c) imaginary parts of the effective refractive index n_{eff} plotted at $E = 1$ eV. The inset in (b) depicts a zoom-in of the region where the spill-out has a significant impact on the mode index. The horizontal dashed gray line marks the refractive index of the bulk metal, n_{bulk}. The dotted red lines represent the region where it becomes hard to find a proper solution due to numerical artifacts. Parts (d)–(f) present a comparison between the electric and magnetic field components considering $g = 1$ nm and $E = 1$ nm. The plots in the right column are normalized with the field value at the center of the gap, except in (e) where the field is zero at the center of the gap and the maximum value of the field is considered for normalization.

down of the mode which is larger than without any form of spill-out. This is understandable, as the Poynting vector inside a plasma is opposite to the propagation direction, leading to a plasmonic drag [46], to a lower group velocity, and finally to a larger wave vector.

When the gap is as small as 0.5 nm it begins to "close," electromagnetically speaking: the electron density inside the gap becomes so large that the propagation of the gap plasmon is hindered. The real part of the effective index decreases quickly, while its imaginary part is much higher when the spill-out is taken into account. It is important to note that in the limit of vanishing gaps, where the spill-out is expected to virtually close the gap, both real and imaginary parts of n_{eff} should ideally converge to the index of the bulk metal. In Fig. 2(c), we notice that when $g < 0.35$ nm, i.e., the gap distances where the gap is expected to vanish completely, the imaginary part of n_{eff} tends to bend back. However, we have found it to be very hard to obtain a proper solution in this region. Indeed, the dotted line as shown in Figs. 2(b) and 2(c) represents a region where we could not find a solution due to numerical artifacts. When $g < 0.28$ nm, we find $n_{eff} = n_{bulk}$, where n_{bulk} is the refractive index of the bulk metal. These results are in agreement with the ones reported in [22].

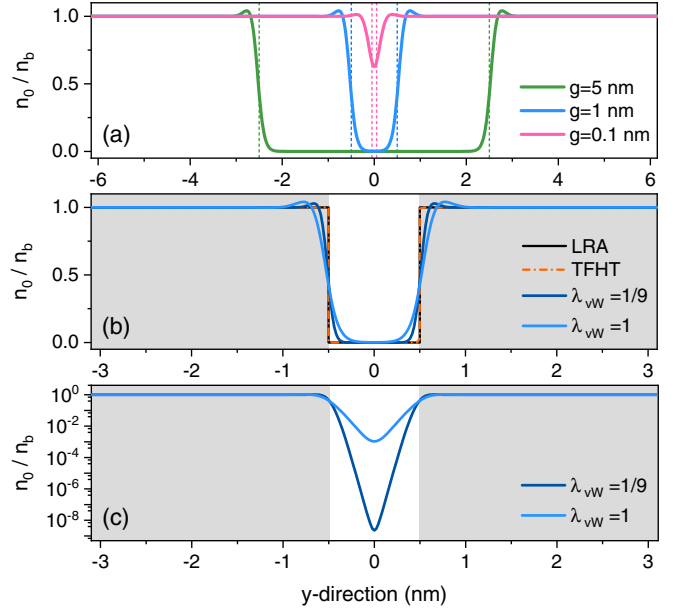We attribute the large imaginary part of the effective index to peaks in the electric field that can be seen when the

FIG. 3. Equilibrium charge density n_0 calculated self-consistently for an air gap sandwiched between the Na metals. (a) For a different gap size (g). The vertical dashed lines represent the metal boundaries for the corresponding gaps. (b) A comparison between the LRA, TFHT, and QHT for the spill-out parameter $\lambda_{vW} = 1/9$ and 1, considering $g = 1$ nm and (c) the decay of the charge density at $\lambda_{vW} = 1/9$ and 1. The results are normalized with the bulk density n_b, and the shaded gray area represents the metal region.

electron density reaches a value so that $\chi_f \simeq 0$, meaning a vanishing local permittivity $\varepsilon(y)$. The components of the electric field in the transverse and longitudinal directions with respect to mode propagation are shown in Figs. 2(d) and 2(e), respectively. The extremely narrow peaks that the electric field exhibits have already been seen in a similar context [22] with a different electron gas density. This means that they are fundamentally linked to the assumption that a local permittivity is able to describe the spill-out accurately. A clue that nonlocality may play an important role here is that $\varepsilon = 0$ is the condition for which volume plasmons are expected to be supported inside the electron gas.

The peaks in the electric field drastically increase the losses experienced by the gap plasmon, since the local absorption by the electron gas is proportional to $\mathbf{E}_1 \cdot \mathbf{J}^*$, where * denotes the complex conjugate [47]. This explains the larger imaginary part. We underline that, on the contrary, the magnetic field remains smooth [see Fig. 2(f)].

B. QHT for simple metals

In the framework of QHT, we compute the self-consistent ground-state density of a symmetric MIM waveguide from Eq. (9). First we consider that the guiding structure is composed of an air gap sandwiched between two semi-infinite Na metals, to better grasp the difference with the simplified spill-out model.

The self-consistent ground-state density $n_0(\mathbf{r})$ for a Na-air-Na configuration is shown in Fig. 3(a) considering different
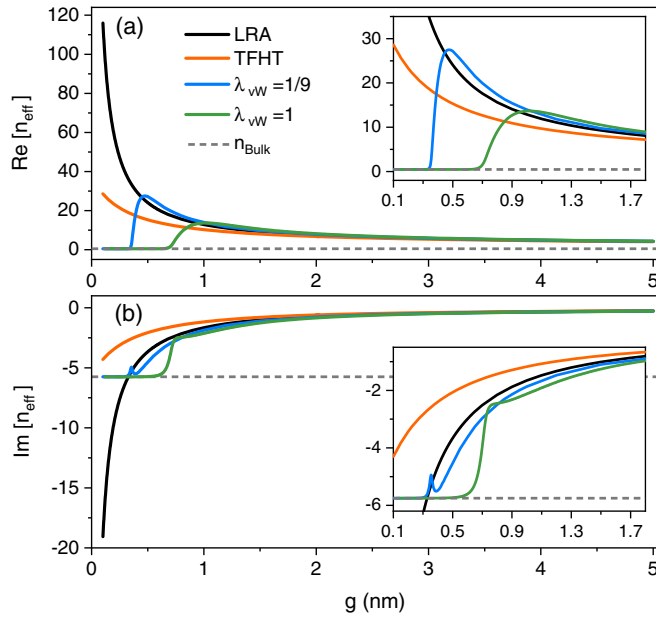
FIG. 4. Effective refractive index n_{eff} of Na-air-Na configuration as a function of gap distance g. A comparison between the LRA, TFHT, and QHT is given. (a) Real and (b) imaginary parts of n_{eff}, plotted at $E = 1$ eV. The horizontal dashed gray lines represent the refractive index of the bulk metal.

FIG. 5. Normalized (a) electric $|E_y|$ and (b) magnetic $|H_x|$ field components plotted along the y-direction, i.e., in the normal direction to the interfaces, plotted under different approaches. The inset of (b) is a closeup view of the $|H_x|$ at the interfaces. The plots are normalized by the field value at the center of the gap. Parts (c) and (d) show decay of the fields at longer distances inside the metal. In these plots, we have considered $g = 1$ nm, $E = 1$ eV, and $\lambda_{\text{vW}} = 1/9$ in the QHT case.

gaps. The metal-dielectric interfaces for each gap size are indicated by the dashed lines. It can be seen that when the interdistance of the metal surfaces is sufficiently large, there is no overlap of the charge densities. In this scenario, the system can be well described using a classical formulation. However, when the gap spacing between the metal boundaries decreases, the charge-density profiles start to overlap and the electrons from the two surfaces start to tunnel across the gap. For sufficiently small gaps, the charge transfer between the metals is large enough to close the gap optically speaking, despite the fact that the metal ion surfaces do not touch physically. Figure 3(b) presents a comparison of n_0 between the LRA, TFHT, and QHT for gap size $g = 1$ nm. In the LRA and the TFHT approximation, we assume a constant $n_0 = n_b$ in the metal and zero outside. In the QHT case, the space-dependent $n_0(\mathbf{r})$ is plotted for both $\lambda_{\text{vW}} = 1/9$ and 1. The parameter $\lambda_{\text{vW}} = 1/9$ represents a lower spill-out and a faster decay of the charge density, whereas $\lambda_{\text{vW}} = 1$ gives a higher spill-out and a slower decay, as shown in Fig. 3(c).

In Fig. 4, we plot the effective refractive index as a function of gap distance g. The results computed within the QHT, considering the self-consistent space-dependent charge density, are compared against the LRA and TFHT without spill-out. It is interesting to note that for extremely small gap separation, the LRA predicts very large values of both real and imaginary parts of n_{eff}, whereas the TFHT gives relatively smaller values. Nonetheless, the trend is quite similar, i.e., n_{eff} increases monotonically as the gap decreases. In contrast, in the QHT approximation, the n_{eff} first increases with the decreasing gap size and then rapidly decreases after a certain gap distance, getting a value of the refractive index of the bulk metal. This is due to the fact that at a certain gap size, a
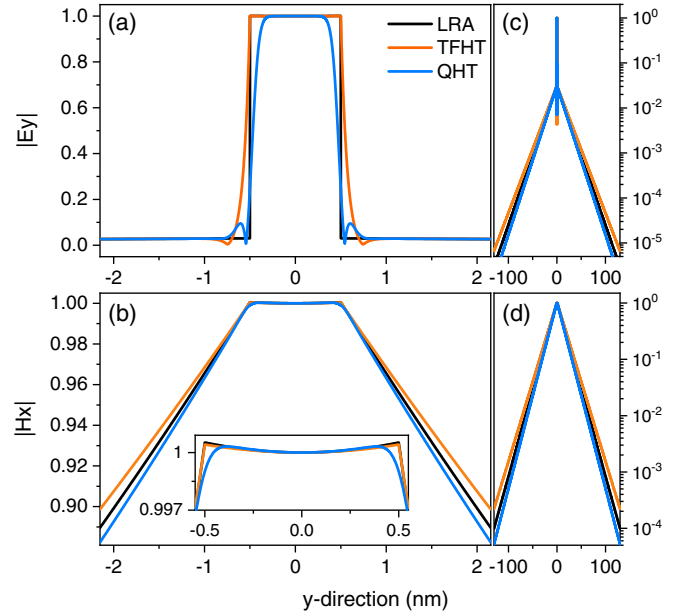
reasonable charge transfer through the gap occurs between the metal surfaces due to the overlapping of the charge densities, and a virtual fade-out of the physical gap appears when the spacing between the metal boundaries is further decreased. For sufficiently small gaps, the mode virtually vanishes due to the electron tunneling, and both the real and imaginary parts of n_{eff} converge to the real and imaginary parts of the refractive index of the bulk metal, as shown in the inset of Figs. 4(a) and 4(b), respectively. The expected limit for the effective index of the gap plasmon is in that case actually retrieved. In the QHT case, results for both $\lambda_{\text{vW}} = 1/9$ and 1 are presented. Since $\lambda_{\text{vW}} = 1/9$ entails the lower spill-out, the gap virtually vanishes when $g < 0.3$ nm, however the gap disappears even at higher separation ($g < 0.45$ nm) for $\lambda_{\text{vW}} = 1$, as it implies higher electron spill-out.

Now, to show the influence of the electron spill-out on the mode profile, we plot here the electric $|E_y|$ and magnetic $|H_x|$ fields in the orthogonal direction to the metal-air interfaces under different approximations for $g = 1$ nm and $E = 1$ eV, as shown in Fig. 5. Both in the LRA and TFHT, $|E_y|$ is constant throughout the gap and then exponentially decays in the metal, as can be seen in Figs. 5(a) and 5(c). QHT, on the other hand, predicts that $|E_y|$ has a maximum value at the center of the gap and it decreases as it moves away from the center. The effect of the electron pressure in the metal in the TFHT can also be observed on the field profile. The $|H_x|$, as plotted in Fig. 5(b), remains quite similar in the gap in all approximations expect near the interfaces, where QHT predicts a bit lower value. The decay of fields is given in Figs. 5(c) and 5(d), which show that the fields decay faster in QHT and slower in TFHT with
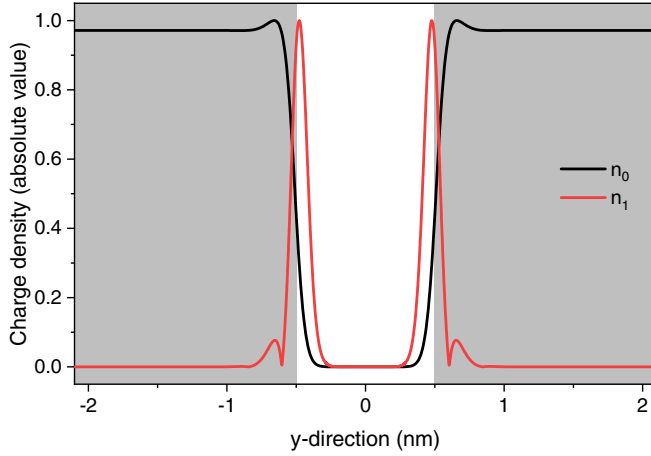
FIG. 6. Normalized equilibrium charge density n_0 and induced charge density n_1 computed using the QHT with $\lambda_{vW} = 1/9$. The other parameters are the same as those used in Fig. 5. The shaded gray area represents the metal regions.

FIG. 7. (a) Real and (b) imaginary parts of n_{eff} within the LRA with electron spill-out and full self-consistent QHT for $\lambda_{vW} = 1/9$. The inset shows a zoom-in of the effective mode index in the tunneling regime. The dotted red lines represent the gap sizes, where it becomes hard to find a proper solution due to numerical artifacts, as was discussed in the caption of Figs. 2(b) and 2(c).

respect to the LRA fields, essentially because the wave vector along z and thus the decay along the y axis is slightly different.

Now, we analyze the impact of electron spill-out and tunneling on the induced charge density. We, in fact, consider the spill-out and tunneling in both the equilibrium charge density and in the linear response, as the modification of the equilibrium density gives a more enhanced overlap of the induced charge densities. The induced charge density indeed is the equivalent of the transition density in the TD-DFT, and, in the limit of the kinetic energy functional approximation, it represents the sum of all possible electron transitions, including those describing the tunneling. As an example, we plot in Fig. 6 the induced charge density n_1 along with the equilibrium charge density n_0 in the same scenario as considered in Fig. 5. The peak of the induced charge density from each interface lies outside the jellium edge, and the tails of the densities go more into the dielectric medium, consequently showing more overlap.

Finally, it is interesting to compare the QHT and the simplified LRA with spill-out, as shown in Fig. 7. The real part of n_{eff} in both cases accords well until the gap closes [Fig. 7(a)], whereas, as described previously, the imaginary part is much higher in the simplified spill-out model [Fig. 7(b)]. This really means that nonlocality avoids the existence of the peaks in the electric field when the electron density reaches a particular value, making the imaginary part much closer to the prediction of the LRA without spill-out.

C. QHT for noble metals

Noble metals, such as Ag, show in general a more complex behavior compared to Drude-type metals due to the contribution of interband transitions and core electrons. In what follows, we use $r_s = 3$ a.u. and $\hbar\gamma = 0.03$ eV for Ag, and we account for interband effects with a constant background dielectric constant $\varepsilon_\infty = 5.9$ [31], which adds a local contribution to the free-electron metal permittivity. The effective refractive index for the Ag-air-Ag configuration as a function of the gap size computed using different models is plotted
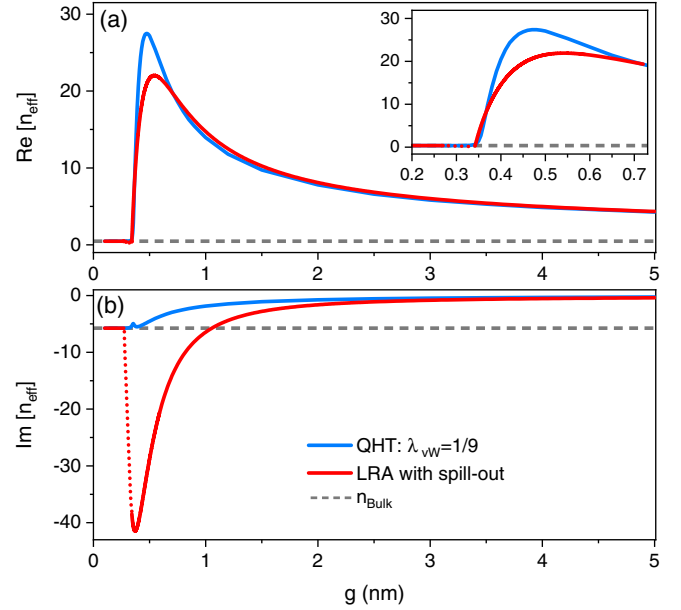
in Fig. 8. The real and imaginary parts of n_{eff} are plotted in Figs. 8(a) and 8(b), respectively, showing quite similar behavior as seen for the Na-Air-Na case in the previous sec-
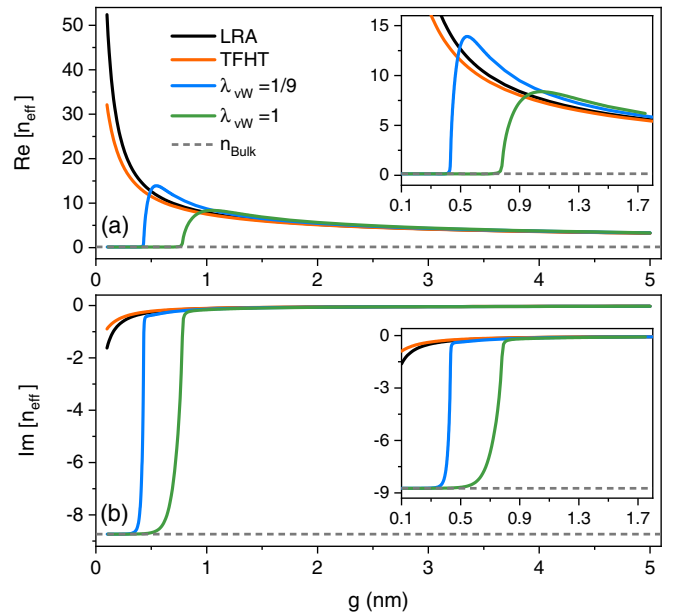

FIG. 8. Effective refractive index for Ag-air-Ag configuration as a function of gap distance. A comparison between the LRA, TFHT, and QHT is given. The upper and lower panels show the real and imaginary parts of the effective index, respectively. The results are plotted at $E = 1$ eV. The dashed gray line represents the refractive index of the bulk metal.
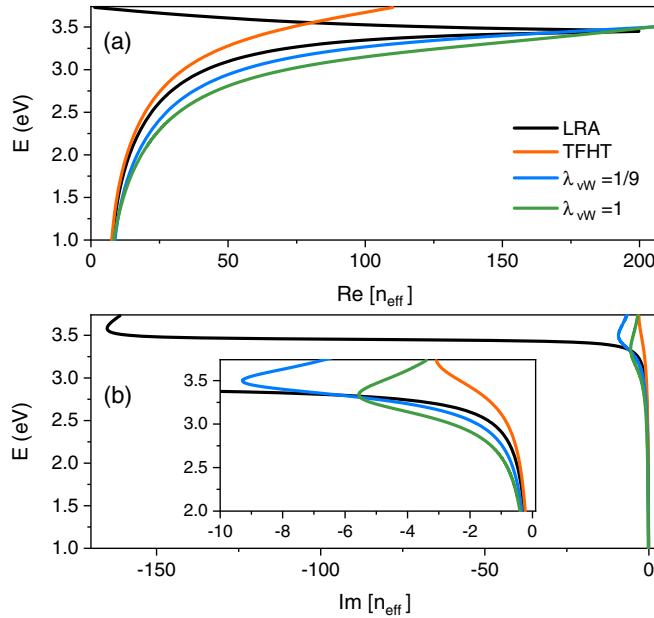
FIG. 9. Effective refractive index as a function of energy (in eV) at $g = 1$ nm. A comparison between the LRA and QHT is given. The upper and lower panels shows the real and imaginary parts of the effective index. The results are normalized with the screened plasma frequency, $\omega'_p = \omega_p/\sqrt{\varepsilon_\infty}$.

tion. LRA and TFHT clearly show the unphysical increasing character of the effective index when subnanometer gaps are in question. QHT, on the other hand, predicts that in the tunneling regime, for sufficiently small metal separation, the gap virtually fades away completely and the effective index converges to the refractive index of the bulk Ag. Note that in Fig. 8, the QHT gives a different threshold for g at which the gap virtually vanishes, depending on the amount of electron spill-out considered from the metal surface, i.e., $\lambda_{vW} = 1/9$ or 1.

So far, we have analyzed the effective mode index of the MIM waveguide as a function of air gap between the metals. It is also important to examine the behavior of n_{eff} as a function of energy. Figures 9(a) and 9(b) present the real and imaginary parts of n_{eff}, respectively, as a function of energy (E in eV) for a Ag-air-Ag waveguide, considering a gap size $g = 1$ nm—a width that is experimentally rather easy to obtain [9].

This underlines the differences between the predictions of the different approaches regarding the effective index of the mode. The LRA predicts a bend back for the real part and a much larger imaginary part in this region. The TFHT, on the contrary, predicts a much lower imaginary part at any frequency. The QHT approach predicts a more reasonable behavior of the gap plasmon with no bend back in the effective index and values of its imaginary part that can be much lower than for the LRA at higher frequencies.

The lower the imaginary part of the effective index, the larger the typical propagation length of the gap plasmon, and finally the larger the quality factor of an eventual cavity for the gap plasmon [48]. This underlines the importance of an accurate prediction in this regime.

IV. EFFECTIVE GAP

To summarize qualitatively the above results, the spill-out seems to be able to close the gap electromagnetically speaking before it is actually spatially or mechanically closed. In addition, we have seen that when the spill-out is taken into account, whether in the simplified model or within the QHT framework, the effective index of the gap plasmon is slightly larger than when the spill-out is neglected, however it presents a very similar trend. All of these results suggest that everything occurs as if the gap were narrower when the spill-out is taken into account, which seems physically sound. The fact that, for the simplified model, peaks seem to appear when the electron density reaches a certain value suggests that it should be possible to define the electromagnetic edges of the metal, as well as an effective value for the gap width. We define this value as the width that should be considered in the LRA model or in the TFHT approach for the gap plasmon to fit the effective index of the gap plasmon provided by the QHT.

The effective gap is thus given by

$$g_e = g - \delta_g,$$

where g is the actual gap between the edges of the metal. Said otherwise, in order to retrieve the results of the QHT, one can use the LRA (or the TFHT) with hard-wall boundaries, but simply use a diminished size of the gap to take the spill-out into account.

For silver, shifting the LRA curve of $\delta_g = 0.15$ nm or the TFHT curve of $\delta_g = 0.22$ nm allows us to retrieve the effective index of the gap plasmon as computed with the QHT extremely accurately, as long as the gap does not start closing (which occurs around $g \simeq 0.6$ nm). The results for the QHT reported in Fig. 10 are for $\lambda_{vW} = 1/9$, however the QHT result for $\lambda_{vW} = 1$ can also be nicely reproduced from the LRA and TF approximations considering $\delta_g \simeq 0.25$ nm and $\delta_g \simeq 0.32$ nm, respectively. The idea that the gap plasmon is sensitive to the effective gap between the metals is thus well founded, as it yields quantitative results. This can be understood because the spill-out causes the response of the electron gas outside the metal to be similar to the response inside the metal, as soon as the electron density is high enough. However, we point out that we have not found another way to define the effective gap based on the profile of the electron density or of the fields. The values we give are simply the ones that yield the best fit, and there does not seem to be any simple way to predict such values.

This discussion finally underlines that at such scales, for the gap plasmon, the tiny difference (a rule of thumb would be, according to our results, to consider a 0.1 nm difference) that exists between the actual edge of the metal and the "electromagnetic edge" has its importance and must be considered in the limit of vanishing gaps.

V. CONCLUSIONS

In this article, we analyzed the impact of both electron spill-out and nonlocality on the propagation of gap plasmons in MIM waveguides using the full self-consistent QHT, which can efficiently take into account both spatial dispersion inside
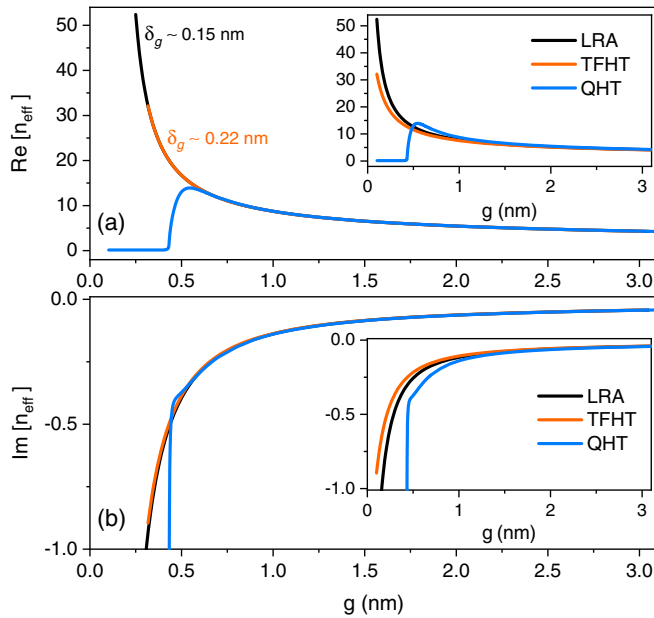
FIG. 10. The effective mode index for Ag-air-Ag waveguide (a) real and (b) imaginary part as a function of gap distance g between the metal edges plotted at $E = 1$ eV. The LRA and TFHT curves are shifted by δ_g to cope with the QHT results, which allows us to introduce the concept of an effective gap that can be defined as $g_e = g - \delta_g$. The g_e seems to appear smaller than the mechanical gap g by an amount of δ_g. In particular, using $\delta_g = 0.15$ nm in the LRA and $\delta_g = 0.22$ nm in the TF approximation gives a good fit with QHT. The insets plot n_{eff} as a function of the actual gap.
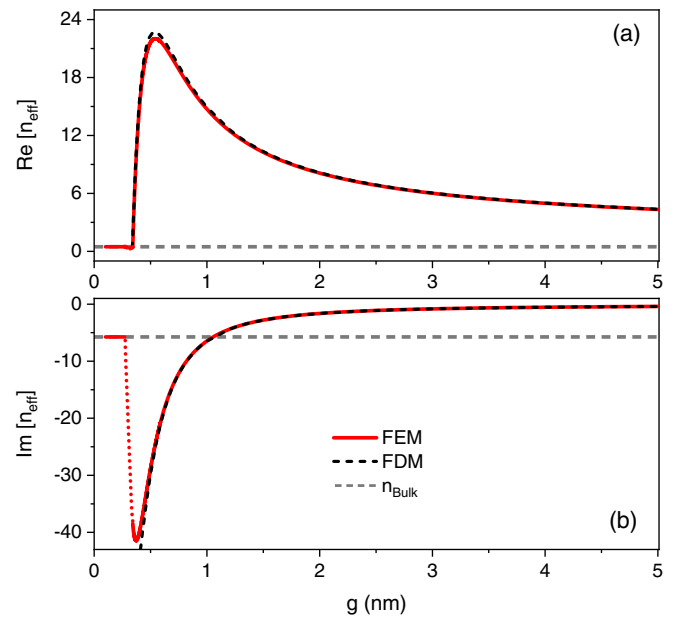

FIG. 11. Comparison between the finite-element method (FEM) and the finite-difference method (FDM). (a) Real and (b) imaginary parts of the effective refractive index n_{eff}. The horizontal dashed gray line represents the refractive index of the bulk Na.

the electron gas and the spatial variation of the electron density outside of the metal. We also compare our results with the commonly used approaches (LRA and TFHT) without any spill-out and with a popular LRA approach for the spill-out [18,22,23].

Peaks in the electric field profile and a large imaginary part for the effective index of the gap plasmon are predicted by the latter approach. Such features are, however, absent in the framework of the QHT, showing that nonlocality is key to an accurate description of the gap plasmon. We underline that only in the QHT framework does the effective mode index converge to the refractive index of the bulk metal, as expected from physical considerations. Hence QHT, without being affected by the artifacts present in the local description, provides a good measure of the effective mode index and of the local field amplitudes, which are particularly relevant for nonlinear optical processes where the response of the system is strongly affected by the local intensities. These points reinforce the idea that QHT is the right tool for plasmonics at the subnanometer level, while using a LRA model for the spill-out may lead, at least in the case of gap-plasmon resonators, to an overestimation of the losses and an underestimation of the quality factor of the resonance entirely due to the fact that nonlocality has been neglected.

Our study thus shows that the spill-out indeed imposes an upper limit on the effective index of the gap plasmon, confirming previous studies [22]. A gap plasmon is very unlikely to present an effective index larger than 15 for noble metals in the

visible, which puts a theoretical limit to the miniaturization of gap-plasmon resonators in the visible.

Importantly, whatever model is chosen to describe the spill-out, it seems possible to introduce an effective gap width between metals, smaller than the gap size defined as the distance between the atoms of the metal across the gap. This effective gap width allows us to take simply the effect of the spill-out into account, as long as the gap can be considered open to the propagation of the mode. In many experiments, the definition of the gap between metals is of crucial importance [7,9]. Our reasoning suggests that using optical characterization methods (like ellipsometry typically) to determine the thickness of a material deposited at the surface of a metal is particularly relevant in the context of gap-plasmon resonators, as the optical thickness that is retrieved probably takes the spill-out partially into account. The gap width deduced from such measurements is in that way probably closer to the effective gap width we have defined here. This may thus explain why hard-wall approaches have so far been surprisingly successful at predicting the experimental behavior of gap plasmons in the limit of vanishing gaps.

APPENDIX: NUMERICAL DETAILS

We implement Eqs. (7), (9), and (10) in a finite-element-method based commercial simulator COMSOL MULTIPHYSICS.

The complete set of equations in their weak formulation, recalling $\partial_t \mathbf{P}_f = \mathbf{J}$, can be expressed as

$$\int (\boldsymbol{\nabla} \times \mathbf{E}) \cdot (\boldsymbol{\nabla} \times \tilde{\mathbf{E}}) - \left(\varepsilon_\infty \frac{\omega^2}{c^2} \mathbf{E} + \mu_0 \omega^2 \mathbf{P}_f \right) \cdot \tilde{\mathbf{E}} \, dV = \mathbf{0},$$
(A1)

$$\int - \frac{e}{m_e} \left(\frac{\delta G[n]}{\delta n} \right) (\boldsymbol{\nabla} \cdot \tilde{\mathbf{P}}_f)$$
$$+ \frac{1}{n_0} \big[(\omega^2 + i\gamma\omega) \mathbf{P}_f + \varepsilon_0 \omega_p^2 \mathbf{E} \big] \cdot \tilde{\mathbf{P}}_f \, dV = 0,$$
(A2)

$$\int -(\boldsymbol{\nabla} \cdot \mathbf{P}_f)(\boldsymbol{\nabla} \cdot \tilde{\mathbf{F}}) - e\mathbf{F} \cdot \tilde{\mathbf{F}} \, dV = \mathbf{0},$$
(A3)

$$\int -\varepsilon_\infty \boldsymbol{\nabla} \left(\frac{\delta G[\xi^2]}{\delta n} \right)_{\xi = \sqrt{n_0}} \cdot \boldsymbol{\nabla}\tilde{\xi} + \frac{e^2}{\varepsilon_0} (\xi^2 - n^+) \tilde{\xi} \, dV = 0,$$
(A4)

where the field quantities with a tilde ($\sim$) indicate the test functions. Since the energy functional contains the second-order derivatives, therefore, we introduce in Eq. (A3) a working variable $\mathbf{F}$, such that $\mathbf{F} = \boldsymbol{\nabla} n_1$, where $n_1 = \frac{1}{e} \boldsymbol{\nabla} \cdot \mathbf{P}_f$. We compute the self-consistent equilibrium charge density n_0 from Eq. (A4), and we find that the solution converges more quickly by using a transformed variable $\xi = \sqrt{n_0}$. Equation (A2) predicts the linear response of the system within the framework of QHT, with $G[n]$ as given by Eq. (20). Assuming $n_0(\mathbf{r}) = n_0$ and $G[n] = T_{\text{TF}}[n]$ with the hard-wall boundary condition gives the standard TFHT, and considering $G[n] = 0$ yields the LRA. In the simplified spill-out model given by Eq. (19), we use the model density as given by Eq. (21), taking electron spill-out into account. We have solved Eq. (19) by an iterative procedure based on a finite-difference scheme. For a thin gap, the decay length of the solution becomes much larger than the gap size. The presence of these two scales is a source of numerical instabilities. To obtain an accurate resolution of the lengthscale associated with the gap, we reduce the computation domain. This is obtained by introducing an artificial interface inside the metal at a distance L from the origin. We solve the problem in the interval $[0, L]$, and we use the analytical solution at the interface as a boundary condition in the finite-difference scheme. To illustrate the numerical scheme, we focus on the term $\partial_y(v(y)\partial_y E_z)$, where we have defined $v \doteq \frac{k_z^2}{\varepsilon(y)k_0^2 - k_z^2}$. The other terms are treated similarly. We discretize the interval $[0, L]$ by an equispaced grid with size Δ. We apply a second-order central scheme

$$\partial_y(v\partial_y E_z)(y_i) = \frac{E_{i-1}v_i^- - E_i(v_i^+ + v_i^-) + E_{i+1}v_i^+}{2\Delta^2},$$

where $E_i = E_z(y_i)$ represents the field at the discrete position $y_i = (i-1)\Delta$ and $v_i^\pm \doteq v(y_{i\pm1}) + v(y_i)$. At the boundary,

$$\partial_y(v\partial_y E_z)(L) = \frac{v(L)\big[E_z(L - \Delta) + E_z(L)(e^{ik_z\Delta} - 2)\big]}{\Delta^2},$$

where we use the fact that for $y > L$, the field has the exponential form $E_z = E_z(L)e^{-ik_z(y-L)}$. We obtain a homogeneous linear system for the discrete field $E(y_i)$ parametrized by the wave number k_z. The plasmon resonances are associated with the k_z for which the system has a nontrivial kernel. We use an iterative procedure based on the Nelder-Mead simplex algorithm [49] where we vary k_z in the complex plane in order to minimize the singular values of the matrix associated with the linear system. The results are depicted in Fig. 11, where we show the agreement of the results obtained by the simplified model based on a finite-difference scheme with the solution of Eqs. (A1)–(A4) obtained by a finite-element method.

[1] J. J. Baumberg, J. Aizpurua, M. H. Mikkelsen, and D. R. Smith, Extreme nanophotonics from ultrathin metallic gaps, Nat. Mater. **18**, 668 (2019).

[2] X. Wang and K. Kempa, Plasmon polaritons in slot waveguides: Simple model calculations and a full nonlocal quantum mechanical treatment, Phys. Rev. B **75**, 245426 (2007).

[3] J. Jung, T. Søndergaard, and S. I. Bozhevolnyi, Gap plasmon-polariton nanoresonators: Scattering enhancement and launching of surface plasmon polaritons, Phys. Rev. B **79**, 035401 (2009).

[4] G. M. Akselrod, C. Argyropoulos, T. B. Hoang, C. Ciracì, C. Fang, J. Huang, D. R. Smith, and M. H. Mikkelsen, Probing the mechanisms of large purcell enhancement in plasmonic nanoantennas, Nat. Photon. **8**, 835 (2014).

[5] T. B. Hoang, G. M. Akselrod, C. Argyropoulos, J. Huang, D. R. Smith, and M. H. Mikkelsen, Ultrafast spontaneous emission source using plasmonic nanoantennas, Nat. Commun. **6**, 7788 (2015).

[6] J. Yang, C. Sauvan, A. Jouanin, S. Collin, J.-L. Pelouard, and P. Lalanne, Ultrasmall metal-insulator-metal nanoresonators: Impact of slow-wave effects on the quality factor, Opt. Express **20**, 16880 (2012).

[7] A. Moreau, C. Ciracì, J. J. Mock, R. T. Hill, Q. Wang, B. J. Wiley, A. Chilkoti, and D. R. Smith, Controlled-reflectance surfaces with film-coupled colloidal nanoantennas, Nature (London) **492**, 86 (2012).

[8] T. Søndergaard, S. M. Novikov, T. Holmgaard, R. L. Eriksen, J. Beermann, Z. Han, K. Pedersen, and S. I. Bozhevolnyi, Plasmonic black gold by adiabatic nanofocusing and absorption of light in ultra-sharp convex grooves, Nat. Commun. **3**, 969 (2012).

[9] C. Ciraci, R. T. Hill, J. J. Mock, Y. Urzhumov, A. I. Fernandez-Dominguez, S. A. Maier, J. B. Pendry, A. Chilkoti, and D. R. Smith, Probing the ultimate limits of plasmonic enhancement, Science **337**, 1072 (2012).

[10] S. Raza, G. Toscano, A.-P. Jauho, M. Wubs, and N. A. Mortensen, Unusual resonances in nanoplasmonic structures due to nonlocal response, Phys. Rev. B **84**, 121412(R) (2011).

[11] C. Ciracì, J. B. Pendry, and D. R. Smith, Hydrodynamic model for plasmonics: A macroscopic approach to a microscopic problem, ChemPhysChem **14**, 1109 (2013).

[12] T. V. Teperik, P. Nordlander, J. Aizpurua, and A. G. Borisov, Robust Subnanometric Plasmon Ruler by Rescaling of the Nonlocal Optical Response, Phys. Rev. Lett. **110**, 263901 (2013).

[13] L. Stella, P. Zhang, F. J. García-Vidal, A. Rubio, and P. García-González, Performance of nonlocal optics when applied to plasmonic nanostructures, J. Phys. Chem. C **117**, 8941 (2013).

[14] W. Zhu, R. Esteban, A. G. Borisov, J. J. Baumberg, P. Nordlander, H. J. Lezec, J. Aizpurua, and K. B. Crozier, Quantum mechanical effects in plasmonic structures with sub-nanometre gaps, Nat. Commun. **7**, 11495 (2016).

[15] J. Zuloaga, E. Prodan, and P. Nordlander, Quantum description of the plasmon resonances of a nanoparticle dimer, Nano Lett. **9**, 887 (2009).

[16] C. A. Ullrich, *Time-Dependent Density-Functional Theory: Concepts and Applications* (Oxford University Press, Oxford, 2011).

[17] G. Aguirregabiria, D. C. Marinica, R. Esteban, A. K. Kazansky, J. Aizpurua, and A. G. Borisov, Role of electron tunneling in the nonlinear response of plasmonic nanogaps, Phys. Rev. B **97**, 115430 (2018).

[18] R. Esteban, A. G. Borisov, P. Nordlander, and J. Aizpurua, Bridging quantum and classical plasmonics with a quantum-corrected model, Nat. Commun. **3**, 825 (2012).

[19] M. N. Feiginov and V. A. Volkov, Self-excitation of 2D plasmons in resonant tunneling diodes, J. Exp. Theor. Phys. Lett. **68**, 662 (1998).

[20] V. Ryzhii and M. Shur, Plasma instability and nonlinear tera-hertz oscillations in resonant-tunneling structures, Jpn. J. Appl. Phys. **40**, 546 (2001).

[21] D. Svintsov, Z. Devizorova, T. Otsuji, and V. Ryzhii, Plasmons in tunnel-coupled graphene layers: Backward waves with quantum cascade gain, Phys. Rev. B **94**, 115301 (2016).

[22] E. J. H. Skjølstrup, T. Søndergaard, and T. G. Pedersen, Quantum spill-out in few-nanometer metal gaps: Effect on gap plasmons and reflectance from ultrasharp groove arrays, Phys. Rev. B **97**, 115429 (2018).

[23] E. J. H. Skjølstrup, T. Søndergaard, and T. G. Pedersen, Quantum spill-out in nanometer-thin gold slabs: Effect on the plasmon mode index and the plasmonic absorption, Phys. Rev. B **99**, 155427 (2019).

[24] G. Toscano, J. Straubel, A. Kwiatkowski, C. Rockstuhl, F. Evers, H. Xu, N. Asger Mortensen, and M. Wubs, Resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmonics, Nat. Commun. **6**, 7132 (2015).

[25] W. Yan, Hydrodynamic theory for quantum plasmonics: Linear-response dynamics of the inhomogeneous electron gas, Phys. Rev. B **91**, 115416 (2015).

[26] C. Ciracì and F. Della Sala, Quantum hydrodynamic theory for plasmonics: Impact of the electron density tail, Phys. Rev. B **93**, 205405 (2016).

[27] C. Ciracì, Current-dependent potential for nonlocal absorption in quantum hydrodynamic theory, Phys. Rev. B **95**, 245434 (2017).

[28] K. Ding and C. T. Chan, Plasmonic modes of polygonal rods calculated using a quantum hydrodynamics method, Phys. Rev. B **96**, 125134 (2017).

[29] S. A. H. Gangaraj and F. Monticone, Do truly unidirectional surface plasmon-polaritons exist? Optica **6**, 1158 (2019).

[30] S. A. Hassani Gangaraj, B. Jin, C. Argyropoulos, and F. Monticone, Broadband Field Enhancement and Giant Nonlinear Effects in Terminated Unidirectional Plasmonic Waveguides, Phys. Rev. Appl. **14**, 054061 (2020).

[31] M. Khalid and C. Ciracì, Enhancing second-harmonic generation with electron spill-out at metallic surfaces, Commun. Phys. **3**, 214 (2020).

[32] E. Runge and E. K. U. Gross, Density-Functional Theory for Time-Dependent Systems, Phys. Rev. Lett. **52**, 997 (1984).

[33] G. Manfredi and F. Haas, Self-consistent fluid model for a quantum electron gas, Phys. Rev. B **64**, 075316 (2001).

[34] N. Crouseilles, P.-A. Hervieux, and G. Manfredi, Quantum hydrodynamic model for the nonlinear electron dynamics in thin metal films, Phys. Rev. B **78**, 155412 (2008).

[35] S. Raza, T. Christensen, M. Wubs, S. I. Bozhevolnyi, and N. A. Mortensen, Nonlocal response in thin-film waveguides: Loss versus nonlocality and breaking of complementarity, Phys. Rev. B **88**, 115401 (2013).

[36] S. I. Bozhevolnyi and J. Jung, Scaling for gap plasmon based waveguides, Opt. Express **16**, 2676 (2008).

[37] C. L. C. Smith, N. Stenger, A. Kristensen, N. A. Mortensen, and S. I. Bozhevolnyi, Gap and channeled plasmons in tapered grooves: A review, Nanoscale **7**, 9355 (2015).

[38] Y. Wang and E. A. Carter, Orbital-free kinetic-energy density functional theory, in *Progress in Theoretical Chemistry and Physics*, edited by S. Schwartz (Kluwer, Dordrecht, 2000), Vol. 5, pp. 117–184.

[39] G. Manfredi, P.-A. Hervieux, and F. Haas, Nonlinear dynamics of electron–positron clusters, New J. Phys. **14**, 075012 (2012).

[40] H. M. Baghramyan, F. Della Sala, and C. Ciracì, Laplacian-Level Quantum Hydrodynamic Theory for Plasmonics, Phys. Rev. X **11**, 011049 (2021).

[41] M. Khalid, F. Della Sala, and C. Ciracì, Optical properties of plasmonic core-shell nanomatryoshkas: A quantum hydrodynamic analysis, Opt. Express **26**, 17322 (2018).

[42] M. Khalid and C. Ciracì, Numerical analysis of nonlocal optical response of metallic nanoshells, Photonics **6**, 39 (2019).

[43] C. Ciracì, R. Jurga, M. Khalid, and F. Della Sala, Plasmonic quantum effects on single-emitter strong coupling, Nanophotonics **8**, 1821 (2019).

[44] F. Yang and K. Ding, Electron spill-out effect in singular metasurfaces, Photonics **8**, 154 (2021).

[45] M. Brack, The physics of simple metal clusters: Self-consistent jellium model and semiclassical approaches, Rev. Mod. Phys. **65**, 677 (1993).

[46] R. Ajib, A. Pitelet, R. Pollès, E. Centeno, Z. Ajaltouni, and A. Moreau, The energy point of view in plasmonics, J. Opt. Soc. Am. B **36**, 1150 (2019).

[47] E. Sakat, A. Moreau, and J.-P. Hugonin, Generalized electromagnetic theorems for nonlocal plasmonics, Phys. Rev. B **103**, 235422 (2021).

[48] C. Lemaître, E. Centeno, and A. Moreau, Interferometric control of the absorption in optical patch antennas, Sci. Rep. **7**, 2941 (2017).

[49] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, Convergence properties of the nelder-mead simplex method in low dimensions., SIAM J. Opt. **9**, 112 (1998).