7-2014

Nonlocal and Quantum-Tunneling Contributions to Harmonic Generation in Nanostructures: Electron-Cloud-Screening Effects

Michael Scalora
Charles M. Bowden Research Center

Maria Antonietta Vincenti
Charles M. Bowden Research Center

Domenico de Ceglia
Charles M. Bowden Research Center

Joseph W. Haus
University of Dayton, jhaus1@udayton.edu

eCommons Citation

Nonlocal and quantum-tunneling contributions to harmonic generation in nanostructures: Electron-cloud-screening effects

Michael Scalora,[1] Maria Antonietta Vincenti,[2] Domenico de Ceglia,[2] and Joseph W. Haus[2,3]

[1]*Charles M. Bowden Research Center, AMRDEC, RDECOM, Redstone Arsenal, Alabama 35898-5000, USA*
[2]*National Research Council–AMRDEC, Charles M. Bowden Research Center, Redstone Arsenal, Alabama 35898, USA*
[3]*Electro-Optics Program, University of Dayton, Dayton, Ohio 45469-2951, USA*

Our theoretical examination of second- and third-harmonic generation from metal-based nanostructures predicts that nonlocal and quantum-tunneling phenomena can significantly exceed expectations based solely on local, classical electromagnetism. Mindful that the diameter of typical transition-metal atoms is approximately 3 Å, we adopt a theoretical model that treats nanometer-size features and/or subnanometer-size gaps or spacers by taking into account (i) the limits imposed by atomic size to fulfill the requirements of continuum electrodynamics, (ii) spillage of the nearly free electron cloud into the surrounding vacuum, and (iii) the increased probability of quantum tunneling as objects are placed in close proximity. Our approach also includes the treatment of bound charges, which add crucial, dynamical components to the dielectric constant that are neglected in the conventional hydrodynamic model, especially in the visible and UV ranges, where interband transitions are important. The model attempts to inject into the classical electrodynamic picture a simple, perhaps more realistic description of the metal surface by incorporating a thin patina of free electrons that screens an internal, polarizable medium.

PACS number(s): 42.65.Ky, 78.67.Uh, 78.68.+m, 71.45.Gm

I. INTRODUCTION

It is generally recognized that theoretical studies of typical optical phenomena that take place at nanometer and sub-nanometer scales necessitate the adoption of methods that go beyond the usual approaches associated with classical electromagnetism. Two relevant examples are nonlocal effects and quantum-tunneling phenomena. Plasmonic phenomena can occur between metallic objects and cavity walls that are in such close proximity that the electronic clouds nearly touch, and an applied electromagnetic field can induce electrons to tunnel between metal objects. Quantum-tunneling phenomena have been addressed using numerically intensive, time-dependent density functional theory (TDDFT), to explore the limitations of classical theory [1–5]. The TDDFT has also been modified into a simpler method referred to as the quantum correction model, which assigns to the gap region the same free-electron properties as the interacting metal components [2]. More recently, in this regard we have developed a quantum conductivity theory, or QCT [6–8], that predicts linear and nonlinear, quantum-induced current densities in the gap region, either a vacuum or a dielectric material, such that the gap itself acquires additional linear and nonlinear optical properties.

While the induced quantum currents tend to limit field enhancement as a result of induced linear and nonlinear absorption, electron tunneling may also facilitate harmonic generation at rates that far exceed typical conversion efficiencies expected for metal nanostructures if quantum tunneling were neglected [9]. For practical purposes, the TDDFT is limited by the number of electronic wave functions that may be used to describe a nanostructure, and so the system under consideration must be small and made of the same metal [1–5,10]. In contrast, the QCT [6–8] generally yields results similar to the TDDFT theory, uses no free parameters, and may be easily combined with Maxwell's equations to explore a wide variety of complex plasmonic systems composed of different metals and insulators, as well as nonlinear optical phenomena that arise as a consequence of quantum tunneling [9].

In addition to quantum tunneling, abrupt changes to the charge density at or near the surface can trigger nonlocal effects. These effects may be studied in a purely classical environment by relating the charge density to the pressure density of an ideal electron gas [11]. The assumption that the electron gas has a quantum nature is separate from the notion of quantum tunneling, and yields a two-component plasma medium whose linear contribution coincides with the classical, ideal gas expression [12,13], and to purely quantum mechanical contributions mostly to harmonic generation if the pump remains undepleted, and to additional nonlocal contributions if the pump energy is drained by a nonlinear conversion process. The result is that the linear dielectric constant turns into a function of frequency and wave vector, i.e., $\varepsilon = \varepsilon(\omega, \mathbf{k})$, and the polarization becomes a function of the field and its spatial derivatives. The effect "softens" the metal surface and smears charges and fields just beneath it. Local [i.e., $\varepsilon = \varepsilon(\omega)$], classical models predict an ever-increasing local field enhancement as the gap between metal components is reduced. In contrast, the inclusion of nonlocal effects in the hydrodynamic model [14–16] typically results in a reduction of the local field in the gap region, accompanied by field penetration that may be exploited to access the metal's nonlinearity [17].

Interest in the study of harmonic generation from metal surfaces has never abated since the early days of nonlinear optics [18–56]. The effective, second-order metal nonlinearity is usually decomposed as separate, tensorial surface and volume contributions [45–56] that have dipolar and quadrupolar origins. Our own, previous treatment of the problem [57,58] was based on extending the hydrodynamic model [19,31,53], which treats conduction electrons only (e.g., 6s shell for Au, 5s shell for Ag), by including explicit, microscopic dipolar [57] and quadrupolar [58] contributions from bound electrons
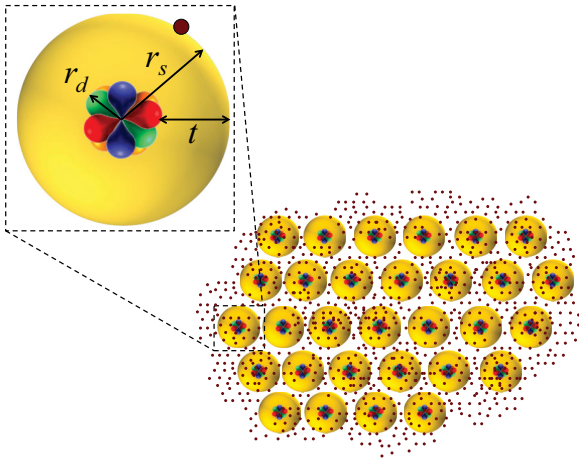
FIG. 1. (Color online) Top left: Scale illustration of d and s orbitals within a single metal atom obtained via Hartree-Fock theory. The radii of the d and s orbitals, r_d and r_s, respectively, correspond to the maxima of each calculated electronic wave function. Parameters for Au: $r_d = 0.64$ Å and $r_s = 1.56$ Å; Ag: $r_d = 0.55$ Å and $r_s = 1.53$ Å; Cu: $r_d = 0.33$ Å and $r_s = 1.37$ Å [61,62], and $t \sim 1$ Å. Bottom right: The electron cloud composed of outer s-shell electrons permeates the entire volume. At the same time, free electrons that belong to atoms near the surface spill outside and screen the hard, ionic background.

(5d shell for Au and 4d shell for Ag), and by making no *a priori* assumptions about what constitutes either surface or volume source. The classical Drude-Lorentz system that we use is well behaved at the surface and yields no unmanageable singularities [51,52]. The inclusion of the linear and nonlinear dynamics of bound electrons may certainly be viewed as an improvement to the hydrodynamic model, especially in near-IR, visible and UV ranges, where the dielectric function deviates significantly from Drude-like behavior. However, an additional, persistent issue is the size of typical surface features and gap or spacer thicknesses, which with modern atomic-layer deposition techniques can easily approach and even be smaller than 1 nm: One has to contend with atomic diameters (or lattice constants) that are of order 3 Å, and an outer-shell electron cloud that may extend several angstroms outside the last atomic surface layer.

Like all classical models, the model exemplified in Refs. [57,58] *does not* contemplate length scales or roughness on the order of the atomic thickness, and *should not* be expected to compare well with purely quantum mechanical approaches [10]. However, the situation may be mitigated by invoking an argument that permits the use of the classical, macroscopic equations in an atomic environment if sources are treated quantum mechanically [59,60]. This approach may be summarized in Fig. 1, where we illustrate scale drawings of a typical transition-metal atom (inset at the top of the figure), complete with inner d shells and outer s orbital. A full-fledged, quantum mechanical Hartree-Fock approach that includes electron-electron interactions and screening may be used to calculate the wave functions associated with each orbital. The wave functions may then be used to deduce orbital radii [61,62]. By orbital radius one generally refers to the distance from the nucleus to the maximum of the wave function (or most likely electron position), which may in fact have several

nodes and be somewhat extended in space [63]. For example, the radius of the uppermost, 5d orbital of gold is approximately $r_d = 0.64$ Å, while the wave function of the 6s-shell peaks at $r_s = 1.56$ Å [61,62], or approximately one atomic radius. Cu and Ag display similar values (see caption of Fig. 1). Then, for atoms arranged in a lattice, the simplest, most rudimentary picture that emerges is similar to the illustration at the bottom of Fig. 1: Nearly free, outer s-shell electrons and bound (inner core) electrons permeate the entire volume, while all rows of atoms near the surface are slightly submerged under s-shell (conduction) electrons (tiny dots in the illustration) that spill outside the metal surface, and in so doing screen the internal medium. Quantum mechanical calculations that assume a uniform, smooth, generic ionic background in fact predict an average, free-electron spill-out distance of approximately 2 Å, which may be understood as roughly the midpoint of a rapidly rarefying medium, i.e., a decaying, electronic wave function whose tail may actually reach somewhat deeper into the surrounding vacuum [59,64–67].

The information contained in Fig. 1 suggests that a classical, Drude-Lorentz oscillator model may be modified to incorporate the basic ideas. One may assume that the medium is composed of an internal, uniform, polarizable mixture of free (Drude) and bound (Lorentz) electrons that extends as far as the outermost reaches of the surface atoms' d orbitals, and by a thin layer of s-shell, free electrons that screens the internal medium. Of course, this is a simplified view that seeks to combine the quantum properties of the atom with macroscopic field equations that are derived in a context where atomic size or roughness must be averaged out, leaving behind only smooth surfaces. In a classical sense, the immediate consequence of the adoption of this physical picture means that the generic, metallic medium that we envision is characterized by a surface layer of finite thickness that has two boundaries: an internal surface, where the (linear and nonlinear) effects of bound charges are extinguished, and an outer surface grazed only by free electrons. For simplicity we assume that the equilibrium (no applied field) charge density of the outer, free-electron shell is constant throughout the layer, although in reality an electronic density gradient is to be expected as the medium becomes more rarefied away from the surface. Density variations of the unilluminated electron cloud as a function of distance from the hard surface could be easily included in the model, with a wave function tail that extends further than 2 Å into the surrounding vacuum, but we expect no significant qualitative or quantitative changes with this added complication.

II. BRIEF OUTLINE OF THE MODEL

The model that we use is based on a microscopic portrayal that begins with a collection of classical Drude-Lorentz oscillators that describe free and bound electrons coupled by the fields. While free charges can move about the entire volume, the motion of bound charges takes place around an equilibrium position that we identify as the radius of the d orbital. This description ultimately gives rise to a patina of free electrons that separates the metal core from the vacuum. The basic equations of motion that couple material and field equations are derived and described in detail in Refs. [57]

and [58] in a context that includes free and bound, dipolar and quadrupolar, and linear and nonlinear sources. Here we provide broad motivation for an approach that includes modification of that model according to the quantization of the atomic structure discussed above, which in turn leads to free-electron screening. We then add the quantum-tunneling parameters for the induced linear and nonlinear currents derived in Ref. [8].

In the absence of quantum tunneling, the generated second-harmonic signal is triggered by free and bound charges alike, because both types of charges interact with the applied fields via intrinsically nonlinear Coulomb (electric) and Lorentz (magnetic) forces. We note that the nonlinear dynamics of bound charges is usually neglected, although it becomes critical at wavelengths below 700 nm, as interband transitions contribute to the dielectric constant. Free charges are also under the action of nonlinear convective and electron gas pressure forces. In addition, the model allows for multipolar, nonlinear source distributions [58] as a result of slight distortions of the inner-core electron cloud resulting from electron screening. In contrast, while a small fraction of the third-harmonic signal always arises from a weak, cascaded process [57] ($3\omega = \omega + 2\omega$, which occurs even if $\chi^{(3)} = 0$) most of it originates from a bulk, third-order nonlinearity attributable to anharmonicities in the motion of bound charges [68].

The effects of the QCT theory are described in detail in Refs. [6–8]. The theory suggests that the gap that separates metal objects fills with induced, linear and nonlinear currents that turn the vacuum or dielectric spacer into an effective medium that displays its own peculiar, linear and nonlinear optical properties. For instance, a vacuum gap approximately $g = 0.8$ nm thick displays an effective $\chi^{(2)} \sim i0.1$ pm/V for adjacent objects composed of dissimilar metals such as Au and Ag, and an effective $\chi^{(3)} \sim i10^{-(20)}$ m²/V² for either similar or dissimilar metals, increasing exponentially for smaller gaps [8,9]. Even though these values may appear to be relatively small, the intensity inside the gap may be amplified thousands of times compared to incident values, thus catalyzing efficient nonlinear optical processes that can far outweigh the intrinsic nonlinearities of the metal.

Our approach thus places free and bound charges on the same footing, adds crucial linear and nonlinear dynamical components to the dielectric constant that are neglected in the conventional hydrodynamic model, treats the full vectorial nature of nonlocal effects without unduly constraining the current normal to the metal surface [14–17,69,70], and also includes the effects of electrons tunneling across the gap. The equations of motion are then integrated in the time domain using a split-step, fast Fourier transform method that propagates the fields, combined with a predictor-corrector method to integrate the material equations [71]. The two-dimensional spatial grid consists of 208×6000 lattice sites discretized in unit cells 1 Å $\times$ 1 Å; the temporal step is 3×10^{-19} s. Reflected and transmitted conversion efficiencies are calculated by sampling the fields at the grid's edges (i.e., far field), and by normalizing the outgoing energy with respect to the total, incident pump energy.

As a final note on the approach, we point out that the model outlined above clearly attempts to account for atomic structure and size, electron spill-out from the surface, and quantum tunneling in order to paint a somewhat more realistic picture of physical phenomena that take place near the metal surface. However, the same may not be said of alternative approaches that introduce artifacts to treat the metal nonlocality [72]. For example, in Ref. [72] the nonlocal metal is replaced outright with a composite material made of local metal covered by a dielectric layer approximately 1 Å thick, (i.e., the approximate Thomas-Fermi length, or in our case, $r_s - r_d$) that even includes the possibility of unphysical gain, ostensibly for the sole reason to ease the computational burden that the metal nonlocality imposes on complex geometrical arrangements. While this may be a clever way to solve the linear problem, the method also injects an arbitrary artifice that upsets the linear and nonlinear postures of surface currents and sources, including nonlinear, nonlocal contributions [53,57], at the same time instilling a false sense of security about the application of effective medium approaches to nonlinear calculations. Then, depending on its precise composition and thickness, the top layer may interfere and perhaps even negate quantum-tunneling effects. Indeed, the modifications that we advocate are easily implemented and pivotal for nonlinear processes like harmonic generation, which may be affected drastically depending on surface properties. Therefore, caution should be exercised when using effective medium models in nonlinear contexts.

III. EXAMPLE CALCULATION FOR A NANOWIRE ARRAY

In Fig. 2 we depict two separate arrays of infinitely long, metal nanowires. Each nanowire is 10 nm in radius (approximately 30 atomic diameters). For both arrays, adjacent cylinders are separated by a distance g, and may be thought of as being composed of either a single metal or dissimilar metals. In Fig. 2(a) we show the way metals are normally treated, if bound charges are considered at all: Free and bound electrons are allowed to be present everywhere, so that d and s orbitals belonging to surface atoms overlap, i.e., $r_b = r_f$, where the
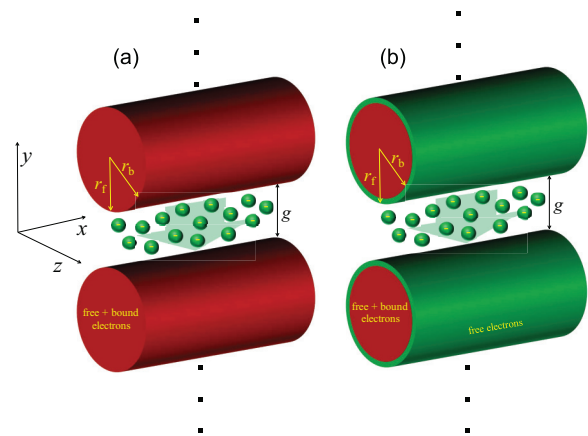

FIG. 2. (Color online) (a) Metal array composed of a mixture of free and bound charges. d and s orbitals overlap, so that bound electron orbits graze the surface. (b) The illustration of Fig. 1 yields a picture where an outer, free-electron shell approximately 2 Å thick covers each nanowire. The green particles inside the gap region of width g represent tunneling electrons.
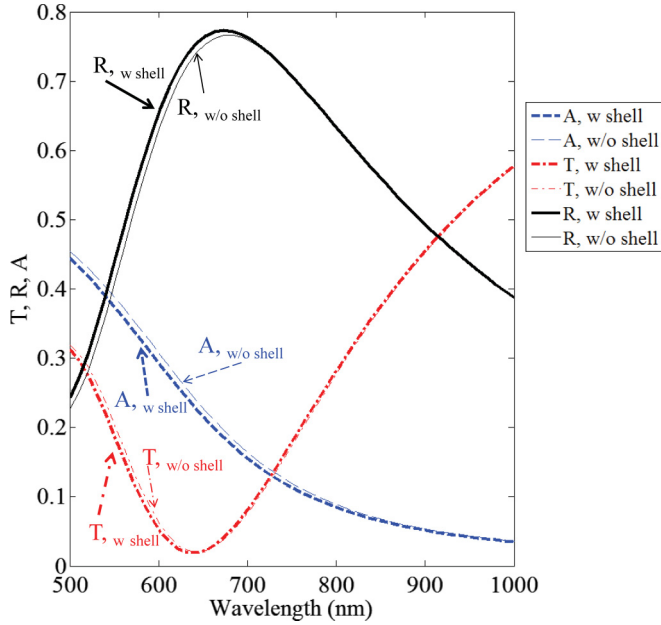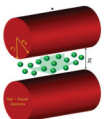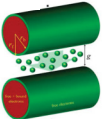
FIG. 3. (Color online) Thin Curves: Reflection, transmission, and absorption vs. wavelength for the gold-silver nanowire array in Fig. 2(a), where free and bound charges extend all the way to the surface. Thick Curves: Reflection, transmission, and absorption for the gold nanowire array in Fig. 2(b), where bound charges are covered by a shallow, free-electron layer.

TABLE I. (Color online) Predicted second-harmonic generation (SHG) and third-harmonic generation (THG) conversion efficiencies.

SH and TH conversion efficiencies Peak Pump Intensity $I_\omega = 1.5$ GW/cm²				
$\eta_{2\omega}$	Local	1.3×10^{-8}	1.8×10^{-9}
	Nonlocal	5×10^{-9}	2.2×10^{-12}
$\eta_{3\omega}$	Local	2.2×10^{-8}	2×10^{-8}
	Nonlocal	4.6×10^{-9}	5×10^{-9}

subscripts b and f stand for bound and free, respectively. In Fig. 2(b) we show the alternative picture that emerges based on Fig. 1, which yields partially overlapping d and s orbitals belonging to surface atoms, so that $r_b < r_f$. The calculated linear transmission, reflection, and absorption spectra for the two scenarios in Fig. 2 are plotted in Fig. 3 for an Au-Ag array, in the local approximation. Nonlocal effects originate in the free-electron gas pressure contribution, lead to decreased local fields, and cause a generic blueshift of the plasmonic band structure, with few additional qualitative or quantitative differences [9], at least in this geometrical arrangement. Palik's gold and silver data [73] are first fitted in the range indicated in the figure using one Drude and one Lorentz oscillator, and are used to calculate the linear spectra of reflection, transmission, and absorption for the array in Fig. 2(a) (thin curves in Fig. 3). The Lorentz component is then removed in a limited region to account for the free-electron-only, green shell shown in Fig. 2(b) and linear spectra are recalculated (thick curves in Fig. 3). The results in Fig. 3 thus show that in the range of interest the *linear* optical properties of the two arrays displayed in Fig. 2 are practically indistinguishable. However, anticipating our nonlinear analysis, the slightly different geometries in Fig. 2 nevertheless lead to large discrepancies in their nonlinear optical properties. The comparison in Fig. 3 thus demonstrates that mere similarities in linear behaviors between full-wave approaches and effective medium models of the kind established in Ref. [72] are not sufficient to also guarantee similarities in nonlinear optical behavior. For simplicity, in what follows we will assume an Au-Ag grating to ultimately excite simultaneously both second- and third-order nonlinearities inside the gap [8,9],

and that the average thickness of the screening, free-electron layer is 2 Å [59,64–67]. We will then compare the results for both types of arrays in order to assess the relevance of the free-electron buffer layer in modeling nanostructures with surface features and gap sizes that approach atomic size, and how nonlocality manifests itself in the two cases, at least according to our model.

In Table I we show the results for the predicted second-harmonic generation (SHG) and third-harmonic generation (THG) conversion efficiencies *without* quantum-tunneling effects, with and without the free-electron buffer layer, with and without nonlocal effects. We assume that both metals exhibit an isotropic, third-order nonlinear response $\chi^{(3)} \approx 10^{-18}$ (m/V)² [68] that reaches the surface in Fig. 2(a) in one case, and is confined to the inner surface in Fig. 2(b) in the other. Pump pulses are 25 fs in duration, are tuned to 700 nm, and have peak power of 1.5 GW/cm². The second-harmonic (SH) signal is tuned to 350 nm and the third-harmonic (TH) signal is tuned to 233 nm. In general, SHG is far more sensitive than THG to surface phenomena because it depends intimately on the evolution and disposition of surface sources, given the centrosymmetric nature of the metal. Without the free-electron screening layer the nonlocal term smears charges and fields away from the surface just enough to reduce the magnitude of the field derivatives (Fig. 4), and hence the amplitudes of nonlinear surface sources, causing a reduction in conversion efficiencies by approximately a factor of 2. In contrast, the introduction of nonlocality when a free-electron buffer layer is present adds elasticity to the surface, voids all surface contributions due to bound charges [58], and reduces conversion efficiencies by nearly three orders of magnitude compared to its local counterpart. If we then compare only *nonlocal* predictions for SHG we find that the free-electron buffer layer suppresses surface contributions from bound charges very effectively and reduces conversion efficiencies by three orders of magnitude. Most of the reduction of SHG conversion efficiency is due to the restriction of bound charges to the inner metal surface: The transition from the inner, red region to the green shell shown in Fig. 2(b) is much smoother compared to the vacuum-metal transition
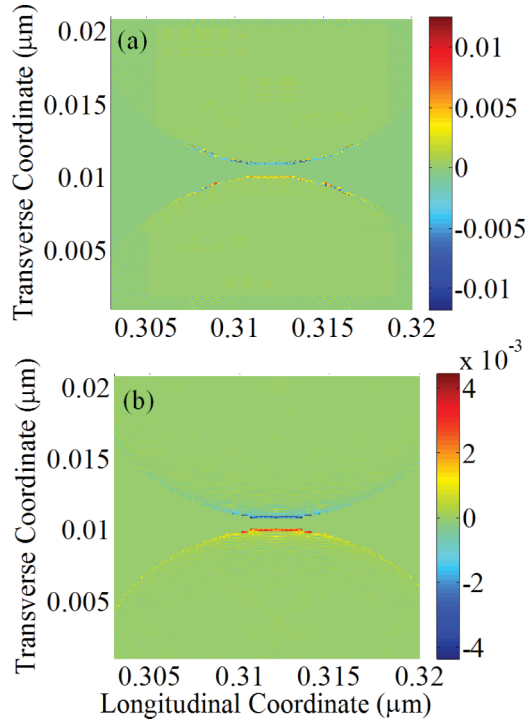
FIG. 4. (Color online) Differential, free charge density distribution δn for (a) local and (b) nonlocal cases. In (b) the metal surface is more elastic compared to (a), and can accommodate a smoother, but lower-amplitude charge distribution around the 2-Å-thick nanoshell. Peak values of δn are larger (by nearly a factor of 4) and sharper in (a) compared to (b); this leads to larger SHG conversion efficiencies in the local case (a) compared to the nonlocal case (b). The distribution in (a) is numerically noisier compared to (b) because the field derivatives are very close to zero, giving rise to unphysical fluctuations inside the volume that have been averaged out.

of Fig. 2(a), which reduces dramatically the influence of nonlinear, bound quadrupolar sources. Our calculations indeed suggest that most of the reduction in SH efficiency is in fact due to the absence of explicit, bound quadrupolar terms. In other words: One structure shields the bound charges, so the bound quadrupolar source does not contribute, hence SHG is low; in the other structure the bound charges are unshielded and bound quadrupoles graze the external surface producing excess SHG at much larger rates. These results should not come as a surprise, since in the wavelength range of interest the dielectric constant is dominated by interband transitions, i.e., the bound electron cloud.

In contrast to SHG, the TH signal *in this particular case* is far less sensitive to the presence of the free-electron buffer layer because the transverse field component couples to the internal, bulk nonlinearity in nearly equal measures in both geometries. Just outside the nanowire, not only is the transverse electric field intensity nearly three orders of magnitude larger than the longitudinally polarized field, it is also shielded far less efficiently. In this case, nonlocal effects reduce conversion efficiencies by nearly a factor of 3 compared to the local case, because the nonlocality reduces overall field amplitudes.

The above observations on THG do not constitute general predictions because slight geometrical changes can strongly

influence the outcome. For example, in Ref. [17] a gold nanowire of square cross section is placed approximately 1 nm above a silver substrate. That arrangement strongly favors the longitudinal component of the field inside the gap region, triggering a localized surface plasmon with an evanescent tail that propels the field into the metal. In that environment, nonlocal effects can either (i) increase THG by nearly three orders of magnitude, if the nanowire has no free-electron buffer layer, as in Fig. 2(a); or (ii) have no influence at all if a free-electron buffer layer only 1 Å thick surrounds the nanowire, which is the approximate spatial separation between d- and s-shell electron orbits, and is sufficient to nearly completely suppress the enhancement of the field normal to the surface [17]. These considerations should serve as further cautionary notes that (i) geometrical considerations always play an important role, and generalization should be avoided; and (ii) the presence of a shielding, free-electron outer layer can dramatically alter predicted SHG and THG conversion efficiencies.

We now illustrate nonlocal effects on charge distribution. In Fig. 4 we show snapshots of the instantaneous, differential free charge density derivable from the continuity equation, in the local and nonlocal approximations, defined as $\delta n = \frac{n(\mathbf{r},t)-n_0}{n_0} = -\frac{1}{en_0}\nabla \cdot \mathbf{P}_f(\mathbf{r},t)$. Although this expression is valid whether or not the nanowires are surrounded by a thin, free-electron-only shell, here we treat the case illustrated in Fig. 2(b). δn is normalized by the equilibrium (no applied field) charge density, n_0, e is the electron charge, and $\mathbf{P}_f(\mathbf{r},t) = \mathbf{P}_\omega(\mathbf{r},t)e^{-i\omega t} + \mathbf{P}_{2\omega}(\mathbf{r},t)e^{-2i\omega t} + \mathbf{P}_{3\omega}(\mathbf{r},t)e^{-3i\omega t} + \text{c.c.}$ is the total, real polarization associated with free charges. If the pump remains undepleted, only the pump term contributes significantly. In Fig. 4(a) δn is calculated in the local approximation; in Fig. 4(b) δn is computed with the addition of the nonlocal electron gas pressure term, as outlined in [57]. We note that in the nonlocal case the maximum amplitude of δn is nearly four times *smaller* compared to its local counterpart, and slightly smeared away from the surface. This smaller, smoother, local surface charge density value explains why the comparison of SHG conversion efficiencies between local and nonlocal cases in Fig. 4(b) strongly favors the local approximation: Nonlinear surface and convective sources are proportional to $(\nabla \cdot \mathbf{P}_f)\mathbf{E}$ and $\dot{\mathbf{P}}_f(\nabla \cdot \mathbf{P}_f) + (\dot{\mathbf{P}}_f \cdot \nabla)\mathbf{P}_f$, respectively [57]; larger spatial derivatives lead to bigger SH conversion efficiencies.

In Fig. 5 we show a snapshot of the spatial distribution of the electric field intensities in the nonlocal case, when the peak of the incident pump pulse reaches the gold-silver grating with the free-electron shell, with and without quantum-induced currents. If adjacent nanowires are made of different metals, and the distance between nanowires is fixed at $g = 0.8$ nm, the currents inside the gap yield a linear dielectric constant $\varepsilon_\omega \approx 1 + i0.4$; $\chi^{(2)} \approx i0.1$ pm/V; and $\chi_\omega^{(3)} \sim \chi_{3\omega}^{(3)} \approx i10^{-20}$ (m/V)^2 [8,9]. We assume incident and generated fields are polarized along the array axis (as shown in Fig. 2). In the quantum-tunneling case excitation of the gap region adds significantly to SHG and THG conversion efficiencies because the pump intensity becomes well localized in the gap, with an enhancement factor of nearly three orders of magnitudes; see Figs. 5(a) and 5(b). However, there are some peculiarities in the field localization and emission properties that we now highlight. At moderate intensities (Fig. 5 corresponds to
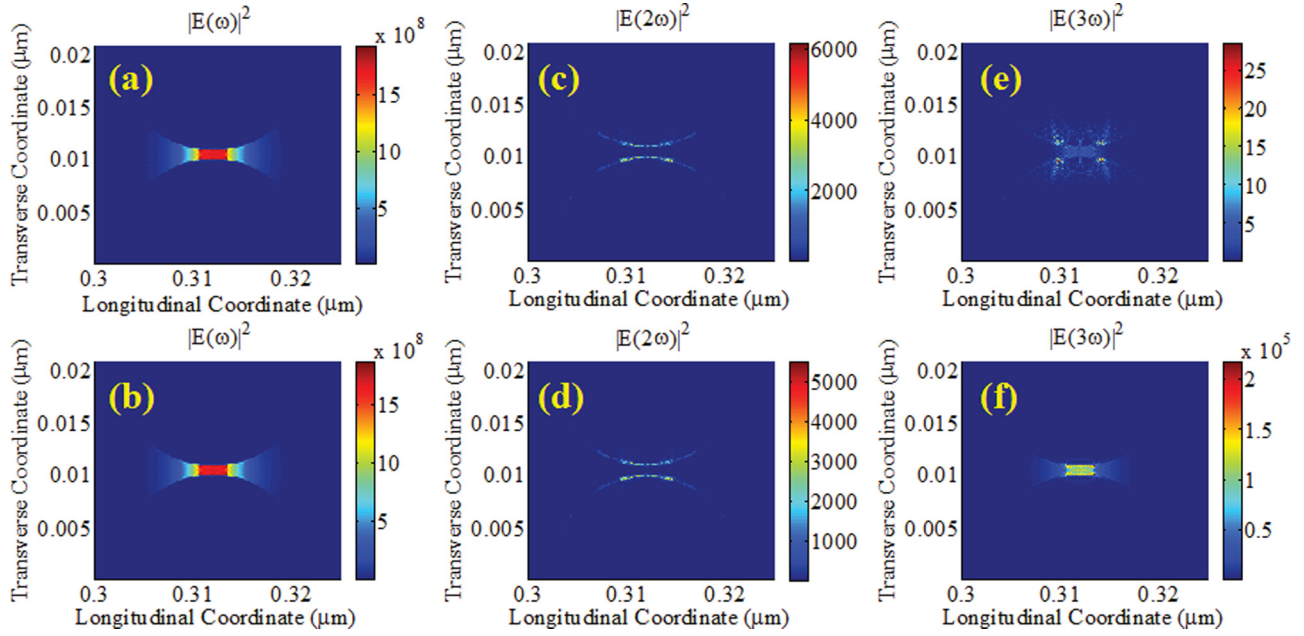
FIG. 5. (Color online) Electric field intensity distributions between two adjacent Au-Ag nanowires without (a), (c), (e) and with (b), (d), (f) quantum-induced linear and nonlinear currents.

0.4 GW/cm^2 peak power), the peak pump field intensity is slightly reduced (less than 3%) by quantum-induced linear absorption. By the same token, in the quantum case the local TH field intensity inside the gap [Fig. 5(f)] is enhanced by nearly four orders of magnitude compared to the classical case [Fig. 5(e)], a quantitative aspect that is also reflected in a corresponding increase in THG conversion efficiency, as reported in Ref. [9]. TH field localization inside the gap is also highly suggestive of the fact that for the most part nonlinear sources are distributed inside the gap, where $\chi_{3\omega}^{(3)} \approx i 10^{-20}$ (m/V)^2, thus overwhelming THG arising from within the metal. However, the SH (near) field localization properties are perhaps the most peculiar. Indeed, quantum-induced currents increase conversion efficiency by nearly three orders of magnitude [9], notwithstanding the fact that the local field intensity decreases by nearly 10% [compare amplitudes in Figs. 5(c) and 5(d)], with field localization properties that are practically unchanged relative to the absence of quantum tunneling. Put another way, a quantum gap is equivalent to a gap doped with a nonlinear material, thus comparable to the introduction of an effective dipolar contribution to the scattered SH light, forbidden in the classical representation. As a consequence, the presence of nonlinear quantum tunneling makes the structure a far more efficient radiator of SH light, even though field localization properties appear to change little compared to the classical case. We believe that this phenomenon is a unique marker of quantum tunneling. While linear effects of quantum tunneling are inherently subtle and hardly distinguishable from nonlocal effects, the nonlinear response is drastically altered by quantum tunneling, especially for SH light, whose nature is forcibly converted from quadrupolar to dipolar.

One needs to bear in mind that material dispersion and incident wavelength are important factors that determine the evolution of the harmonic fields. In general, in the wavelength range under consideration (below 700 nm) the linear dielectric response of the metal at the harmonic wavelengths (deep in the UV range) is dominated by the dynamics of bound electrons. However, the nonlinear response in fact also appears to be regulated by the presence of a screening, free-electron layer, which becomes one of the dominant features in harmonic generation. As an example, pumping the grating at 600 nm increases THG (at 200 nm) by two orders of magnitude compared to pumping at 700 nm, thanks to a combination of improved resonance conditions and field penetration depth, and reduced Im(ε) for both Ag and Au.

Finally, we note that according to the QCT model [6–9] the magnitude of the quantum-induced coefficients increases at near-exponential rates for decreasing gap sizes. For example, according to the model a gap $g = 0.6$ nm wide will display increased linear absorption and nearly two orders of magnitude enhancement in the nonlinear coefficients compared to a 0.8-nm gap. By the same token, a different geometrical arrangement may offer far improved field localization characteristics, as in Ref. [17], for example, where the local field intensity is enhanced more than 10^4 times, which may suffice to trigger quantum-tunneling events for slightly larger gap sizes [8].

IV. CONCLUSIONS

We have presented a theoretical model that allows the study of linear and nonlinear optical phenomena like SHG and THG from nanoplasmonic environments in a context that takes into account (i) linear and nonlinear dynamics of the bound electron cloud; (ii) electron spill-out effects and resultant screening of an internal, polarizable medium; and (iii) electronic quantum-tunneling effects that induce linear and nonlinear currents between two metal objects placed in close proximity. We have investigated harmonic generation and

compared the results in the local and nonlocal approximations, with and without electron screening conditions. In the absence of quantum tunneling, which for vacuum translates to gap sizes of order 1 nm, and up to ~ 2 nm for appropriate dielectric materials [8], our results suggest that both SHG and THG are sensitive to the geometry, the screening effects of a free-electron cloud that surrounds the internal medium, and nonlocal effects. For sufficiently small gap sizes, harmonic generation originating inside the quantum gap can easily overcome, by several orders of magnitude, the amplitudes of the harmonic signals arising from the intrinsic nonlinearities of the metal [9]. Our results thus suggest that both quadratic and cubic nonlinear effects may be triggered by quantum tunneling and may be observed in the far field at practical light irradiance levels (~ 1 GW/cm^2), using femtosecond pulses,

from metallic nanostructures with gap sizes of the order of 1 nm. Linear optical experiments designed with exquisite, atomic-layer by atomic-layer control over the spacing between the structures were reported in [16,74,75]. New experiments may be designed in similar fashion to study the nonlinear response.

[1] J. Zuloaga, E. Prodan, and P. Nordlander, Nano Lett. **9**, 887 (2009).

[2] J. Zuloaga, E. Prodan, and P. Nordlander, ACS Nano **4**, 5269 (2010).

[3] D. C. Marinica, A. K. Kazansky, P. Nordlander, J. Aizpurua, and A. G. Borisov, Nano Lett. **12**, 1333 (2012).

[4] R. Esteban, A. G. Borisov, P. Nordlander, and J. Aizpurua, Nat. Commun. **3**, 825 (2012).

[5] T. V. Teperik, P. Nordlander, J. Aizpurua, and A. G. Borisov, Opt. Express **21**, 27306 (2013).

[6] J. W. Haus, L. Li, N. Katte, C. Deng, M. Scalora, D. de Ceglia, and M. A. Vincenti, Proc. SPIE **8883**, 888303 (2013).

[7] J. W. Haus, D. de Ceglia, M. A. Vincenti, and M. Scalora, J. Opt. Soc. Am. B **31**, 259 (2014)

[8] J. W. Haus, D. de Ceglia, M. A. Vincenti, and M. Scalora, Proc. SPIE **8994**, 89941Q (2014).

[9] J. W. Haus, D. de Ceglia, M. A. Vincenti, and M. Scalora, J. Opt. Soc. Am. B **31**, A13 (2014).

[10] V. Kulkarni, E. Prodan, and P. Nordlander, Nano Lett. **13**, 5873 (2013).

[11] E. L. Linder, Phys. Rev. **49**, 753 (1936).

[12] G. Manfredi and F. Haas, Phys. Rev. B **64**, 075316 (2001).

[13] N. Crouseilles, P. A. Hervieux, and G. Manfredi, Phys. Rev. B **78**, 155412 (2008).

[14] F. J. Garcia de Abajo, J. Phys. Chem. B **112**, 17983 (2008).

[15] J. M. McMahon, S. K. Gray, and G. C. Schatz, Nano Lett. **10**, 3473 (2010).

[16] 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, Science **337**, 1072 (2012).

[17] C. Ciracì, M. Scalora, and D. R. Smith, arXiv:1311.3194.

[18] E. Adler, Phys. Rev. **134**, A728 (1964).

[19] N. Bloembergen, R. K. Chang, S. S. Jha, and C. H. Lee, Phys. Rev. **174**, 813 (1968).

[20] S. Jha, Phys. Rev. Lett. **15**, 412 (1965).

[21] S. Jha, Phys. Rev. **140**, A2020 (1965).

[22] F. Brown, R. E. Parks, and A. M. Sleeper, Phys. Rev. Lett. **14**, 1029 (1965).

[23] F. Brown and R. E. Parks, Phys. Rev. Lett. **16**, 507 (1966).

[24] N. Bloembergen and Y. R. Shen, Phys. Rev. **141**, 298 (1966).

[25] S. Jha and C. S. Warke, Phys. Rev. **153**, 751 (1967).

[26] N. Bloembergen and C. H. Lee, Phys. Rev. Lett. **19**, 835 (1967).

[27] H. Sonnenberg and H. Heffner, J. Opt. Soc. Am. **58**, 209 (1968).

[28] J. Rudnick and E. A. Stern, Phys. Rev. B **4**, 4274 (1971).

[29] W. K. Burns and N. Bloembergen, Phys. Rev. B **4**, 3437 (1971).

[30] J. C. Quail and H. J. Simon, Phys. Rev. B **31**, 4900 (1985).

[31] M. Corvi and W. L. Schaich, Phys. Rev. B **33**, 3688 (1986).

[32] N. N. Akhmediev, I. V. Mel'nikov, and L. J. Robur, Laser Phys. **4**, 1194 (1994).

[33] K. A. O'Donnell and R. Torre, New J. Phys. **7**, 154 (2005).

[34] M. Airola, Y. Liu, and S. Blair, J. Opt. A: Pure Appl. Opt. **7**, S118 (2005).

[35] J. A. H. van Nieuwstadt, M. Sandtke, R. H. Harmsen, F. B. Segerink, J. C. Prangsma, S. Enoch, and L. Kuipers, Phys. Rev. Lett. **97**, 146102 (2006).

[36] M. W. Klein, M. Wegener, N. Feth, and S. Linden, Opt. Express **15**, 5238 (2007).

[37] M. C. Larciprete, A. Belardini, M. G. Cappeddu, D. de Ceglia, M. Centini, E. Fazio, C. Sibilia, M. J. Bloemer, and M. Scalora, Phys. Rev. A **77**, 013809 (2008).

[38] Y. Zeng, W. Hoyer, J. Liu, S. W. Koch, and J. V. Moloney, Phys. Rev. B **79**, 235109 (2009).

[39] V. K. Valev, N. Smisdom, A. V. Silhanek, B. De Clercq, W. Gillijns, M. Ameloot, V. V. Moshchalkov, and T. Verbiest, Nano Lett. **9**, 3945 (2009).

[40] T. Xu, X. Jiao, and S. Blair, Opt. Express **17**, 23582 (2009).

[41] V. K. Valev, A. V. Silhanek, N. Smisdom, B. De Clercq, W. Gillijns, O. A. Aktsipetrov, M. Ameloot, V. V. Moshchalkov, and T. Verbiest, Opt. Express **18**, 8286 (2010).

[42] J. Butet, J. Duboisset, G. Bachelier, I. Russier-Antoine, E. Benichou, C. Jonin, and P. F. Brevet, Nano Lett. **10**, 1717 (2010).

[43] V. K. Valev, A. V. Silhanek, N. Verellen, W. Gillijns, P. VanDorpe, O. A. Aktsipetrov, G. A. E. Vandenbosch, V. V. Moshchalkov, and T. Verbiest, Phys. Rev. Lett. **104**, 127401 (2010).

[44] S. Palomba, H. Harutyunyan, J. Renger, R. Quidant, N. F. van Hulst, and L. Novotny, Philos. Trans. R. Soc., A **369**, 3497 (2011).

[45] W. L. Schaich and B. S. Mendoza, Phys. Rev. B **45**, 14279 (1992).

[46] C. D. Hu, Phys. Rev. B **40**, 7520 (1989).

[47] A. Liebsch, Phys. Rev. Lett. **61**, 1233 (1988).

[48] D. Krause, C. W. Teplin, and C. T. Rogers, J. Appl. Phys. **96**, 3626 (2004).

[49] F. X. Wang, F. J. Rodríguez, W. M. Albers, R. Ahorinta, J. E. Sipe, and M. Kauranen, Phys. Rev. B **80**, 233402 (2009).

[50] Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 2002).

[51] D. Maystre, M. Neviere, and R. Reinish, Appl. Phys. A **39**, 115 (1986).

[52] J. L. Coutaz, D. Maystre, M. Neviere, and R. Reinisch, J. Appl. Phys. **62**, 1529 (1987).

[53] J. E. Sipe, V. C. Y. So, M. Fukui, and G. I. Stegeman, Phys. Rev. B **21**, 4389 (1980); J. E. Sipe and G. I. Stegeman, in *Surface Polaritons: Electromagnetic Waves at Surfaces and Interfaces*, edited by V. M. Agranovich and D. Mills (North-Holland, Amsterdam, 1982).

[54] A. Benedetti, M. Centini, C. Sibilia, and M. Bertolotti, J. Opt. Soc. Am. B **27**, 408 (2010).

[55] A. Benedetti, M. Centini, M. Bertolotti, and C. Sibilia, Opt. Express **19**, 26752 (2011).

[56] M. Centini, A. Benedetti, C. Sibilia, and M. Bertolotti, Opt. Express **19**, 8218 (2011).

[57] M. Scalora, M. A. Vincenti, D. de Ceglia, V. Roppo, M. Centini, N. Akozbek, and M. J. Bloemer, Phys. Rev. A **82**, 043828 (2010).

[58] M. Scalora, M. A. Vincenti, D. de Ceglia, N. Akozbek, V. Roppo, M. J. Bloemer, and J. W. Haus, Phys. Rev. A **85**, 053809 (2012).

[59] J. D. Jackson, *The Classical Electromagnetic Field* (Wiley, New York, 1999).

[60] L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (Dover, New York, 1987).

[61] J. B. Mann, *Atomic Structure Calculations II. Hartree-Fock Wave Functions and Radial expectation Values: Hydrogen to Lawrencium*, Report LA-3691 (Los Alamos Scientific Laboratory, Los Alamos, NM, 1968).

[62] http://www.webelements.com/.

[63] M. Kaupp, J. Comput. Chem. **28**, 320 (2007).

[64] N. D. Lang and W. Kohn, Phys. Rev. B **1**, 4555 (1970).

[65] N. D. Lang and W. Kohn, Phys. Rev. B **3**, 1215 (1971).

[66] V. E. Kenner, R. E. Allen, and W. M. Saslow, Phys. Lett. A **38**, 255 (1972).

[67] A. Liebsch, Phys. Rev. B **48**, 11317 (1993).

[68] R. W. Boyd, *Nonlinear Optics* (Academic, Burlington, MA, 2003).

[69] N. A. Mortensen, S. Raza, M. Wubs, T. Søndergaard, and S. I. Bozhevolnyi, Nat. Commun. **5**, 3809 (2014).

[70] S. Raza, M. Wubs, S. I. Bozhevolnyi, and N. A. Mortensen, arXiv:1406.5091.

[71] M. Scalora, M. A. Vincenti, D. de Ceglia, M. Grande, and J. W. Haus, J. Opt. Soc. Am. B **29**, 2035 (2012).

[72] Y. Luo, A. I. Fernandez-Dominguez, A. Wiener, S. A. Maier, and J. B. Pendry, Phys. Rev. Lett. **111**, 093901 (2013).

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

[74] K. J. Savage, M. M. Hawkeye, R. Esteban, A. G. Borisov, J. Aizpurua, and J. J. Baumberg, Nature **491**, 574 (2012).

[75] J. A. Scholl, A. García-Etxarri, A. L. Koh, and J. A. Dionne, Nano Lett. **13**, 564 (2013).