Second- and third-harmonic generation in metal-based structures

M. Scalora,[1] M. A. Vincenti,[2] D. de Ceglia,[2] V. Roppo,[3] M. Centini,[4] N. Akozbek,[1] and M. J. Bloemer[1]
[1]*Charles M. Bowden Research Center AMSRD-AMR-WS-ST, RDECOM, Redstone Arsenal, Alabama 35898-5000, USA*
[2]*AEgis Technologies Group, 410 Jan Davis Drive, Huntsville, Alabama 35806, USA*
[3]*Universitat Politècnica de Catalunya, Departament de Física i Enginyeria Nuclear, Colom 11, E-08222 Terrassa, Spain*
[4]*Dipartimento di Energetica, University of Rome La Sapienza, Via Scarpa 16, I-00185 Rome, Italy*
(Received 14 June 2010; published 18 October 2010)

We present a theoretical approach to the study of second- and third-harmonic generation from metallic structures and nanocavities filled with a nonlinear material in the ultrashort pulse regime. We model the metal as a two-component medium, using the hydrodynamic model to describe free electrons and Lorentz oscillators to account for core electron contributions to both the linear dielectric constant and harmonic generation. The active nonlinear medium that may fill a metallic nanocavity, or be positioned between metallic layers in a stack, is also modeled using Lorentz oscillators and surface phenomena due to symmetry breaking are taken into account. We study the effects of incident TE- and TM-polarized fields and show that a simple reexamination of the basic equations reveals additional, exploitable dynamical features of nonlinear frequency conversion in plasmonic nanostructures.

DOI: 10.1103/PhysRevA.82.043828 PACS number(s): 42.65.Ky, 42.70.Qs

I. INTRODUCTION

Interest in nonlinear frequency conversion in metals and semiconductors alike arches back to the beginning of nonlinear optics [1–15]. Recent research in linear plasmonic phenomena like subwavelength resolution [16] and enhanced transmission [17] has focused renewed attention on the origins of harmonic generation in metamaterials [18–23]; metallic substrates with empty holes [24–27]; holes filled with GaAs [28]; resonant, subwavelength nanocavities [29]; and layered metal-dielectric photonic band gap structures [30,31]. These studies have shown that generation and enhancement of harmonic frequencies are possible in a variety of conditions and circumstances.

In the case of harmonic generation in metals it is notoriously difficult to reconcile quantitative and qualitative aspects of theory and experiments simultaneously. Usually, experimental results can be explained qualitatively by separating the nonlinear contributions into surface and volume sources and by assigning to them suitable weights [32–39]. Our aim here is to study the dynamics in the ultrashort pulse regime, with an eye toward achieving as much qualitative and quantitative agreement as possible between theory and experiments, without imposing any separation between surface and volume sources. We treat free electrons using the hydrodynamic model [3,40–42], make no *a priori* assumptions about charge or current distributions, and include Coulomb, Lorentz, convective, electron gas pressure, and linear and nonlinear contributions to the linear dielectric constant of the metal arising from bound (or valence) electrons.

It has been shown that contributions to second-harmonic generation (SHG) from bound charges can be significant [10]. Free and bound electrons act in similar ways by displaying surface and volume sources, so the general form of the nonlinear source may be specified in terms of a complex dielectric function defined at the fundamental and the SH frequencies [7,11]. Bound electrons contribute to the linear dielectric constant of typical metals (interband transitions) at near-IR wavelengths for gold [43] and copper, with more pronounced effects in the visible and UV ranges [11,44]. Even silver [45] departs from a simple Drude description at near-IR wavelengths. Its linear dielectric function is adequately described by a combined Drude-Lorentz model that contains a mix of free and bound electrons having one or more resonances at UV wavelengths [46–48]. That is:

$$\varepsilon(\omega) = 1 - \frac{\tilde{\omega}_{pf}^2}{\omega^2 + i\tilde{\gamma}_f\omega} - \frac{\tilde{\omega}_{pb}^2}{\omega^2 - \tilde{\omega}_{0,b}^2 + i\tilde{\gamma}_b\omega}. \qquad (1)$$

All parameters are scaled in units of μm^{-1}. $\tilde{\omega}_{p,f}$ and $\tilde{\gamma}_f$ are the plasma frequency and damping coefficient for free electrons; $\tilde{\omega}_{p,b}, \tilde{\gamma}_b$, and $\tilde{\omega}_{0,b}$ are the plasma frequency, damping coefficient, and resonance frequency for bound electrons. By choosing $(\tilde{\omega}_{p,f}, \tilde{\gamma}_f) = (0.0573, 6.965)$ and $(\tilde{\omega}_{p,b}, \tilde{\gamma}_b, \tilde{\omega}_{0,b}) = (0.526, 3.96, 3.1)$ Eq. (1) becomes a fairly accurate representation of the dielectric function down to approximately 400 nm. In Fig. 1 we compare the data found in Ref. [45] with Eq. (1). The solid black curve that runs below all others is the $\text{Re}(\varepsilon)$ of only the Drude portion of Eq. (1). The figure shows that it is not possible to fit any two points on the real portion of the data curve using only a Drude function and a single plasma frequency, highlighting the importance of bound electrons. In this vein, in Ref. [30] the dynamics of the fundamental (800 nm) and SH (400 nm) fields was modeled using free electrons only. That kind of approach forces the use of two distinct free-electron plasma frequencies for each field, even at wavelengths where the impact of core electrons is considerable. This modeling practice leads to the neglect of the dynamics of core electrons. Instead, the mere introduction of bound electrons can recalibrate the linear dielectric function (phase) and its slope (group or energy velocity), which for computational purposes is akin to modifying effective plasma frequency, electron mass, and density [46]. Later in the examples we will see that small changes in the effective electron mass and density in conduction and valence bands can lead to notable qualitative and quantitative differences in the results.

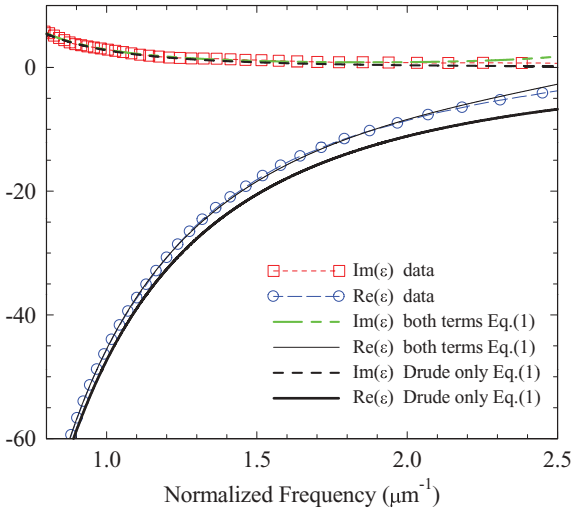

FIG. 1. (Color online) Markers: Re(ε) and Im(ε) parts for Ag in the 1064- to 400-nm range. The solid, thin black, and dashed green curves correspond to the full Eq. (1) and retrace the data well. The lower thick black curve is only the Re(ε) of the Drude portion of Eq. (1). At 450 nm (normalized frequency of 2.5) the discrepancy with the data is substantial. The Im(ε) from just the Drude part in Eq. (1) (dashed black curve) still fits the data reasonably well.

In contexts similar to those of Ref. [28], where a metallic nanocavity is filled with GaAs, there is a tendency to focus only on nonlinear restoring forces, to always neglect intrinsically nonlinear magnetic forces that drive all bound electrons, and to ignore harmonic generation arising from the metal itself. Indeed, while magnetic forces in bound electrons may be several orders of magnitude smaller than nonlinear restoring forces, they are always present nonetheless and in fact play a catalytic role by activating new interaction channels. The nonlinear frequency conversion of GaAs-filled holes on a gold substrate has been investigated at 3 μm [28], with reported TE-polarized SH (1.5-μm) conversion efficiencies comparable to a standard quasi-phase-matched lithium niobate sample of similar thickness. In Ref. [29] it was reported that it is possible to calibrate the width of even a single aperture carved on a silver substrate of specified thickness in order to achieve enhanced transmission, field localization, and, as a consequence, enhanced second-harmonic generation that is strongly correlated to linear transmission maxima. The subwavelength cavity thus designed is capable of localizing and enhancing the incident fields by several orders of magnitude, with effects such as energy velocity reduction down to values less than $c/100$. It is therefore possible to amplify the nonlinear response of a nanocavity filled with a nonlinear material far more than reported in Ref. [28], for both TE- and TM-polarized harmonics, provided the cavity is aptly designed. These results will be reported separately.

II. METALS

Even if metals do not have intrinsic quadratic nonlinear terms, many early works consistently reported on SHG based on the mere presence of the Lorentz force induced by the incident magnetic field. Using a classical oscillator electron model, SH source terms consisting of a magnetic dipole due to the Lorentz force and an electric quadrupolelike contribution emerge also in centrosymmetric media [4–7]. Subsequent experimental work confirmed the existence of at least two SH source terms [8,9] (volume and surface contributions) that may be excited by a polarization normal or parallel to the plane of incidence, respectively. Even though this separation may be practical, it is generally recognized that it may not be possible to fully decompose nonlinear sources as surface and volume terms due at least in part to the presence of spatial derivatives on the bulk polarization [49]. These terms introduce some arbitrariness in the effective surface and volume coefficients, making a clear distinction between these two types of sources almost impossible to carry out [36].

The situation is more ambiguous for metal layers only a few tens of nanometers thick, where field penetration and localization occur inside the metal itself [30]. One might inquire about the relative importance of convective versus Lorentz or Coulomb terms [19], how the presence of sharp corners in metallic nanocavities [29] or nanotubes [50] changes the relative contributions with respect to each other, or what role bound charges and electron gas pressure terms play. What follows is an attempt to address these issues and to highlight as much as possible the most salient dynamical aspects of the interaction of free and bound charges in the ultrashort pulse regime.

It is well known that SHG can be enhanced by coupling with surface plasmons [15,40]: the signal generated by reflection of the incident beam from a metal surface has important contributions from currents stimulated near the surface. In an effort to consider all the forces that act on the electrons as fully as possible in the ultrafast regime we have developed a detailed analysis of SHG and third-harmonic generation (THG) from metallic surfaces and nanocavities that may contain a nonlinear material. The metal is composed of free electrons that occupy the conduction band (typically one electron per atom; for Ag the uppermost level is the $5s^1$) and electrons that fill the valence band (for Ag the uppermost filled valence level is the $4d^{10}$ orbital, with ten available electrons). Free electrons are described by the hydrodynamic model [3,40–42,51], while ordinary Lorentz oscillators are used to describe bound charges. Although this is a simplified picture of metals, in what follows we derive and integrate equations of motion that couple free and bound electrons in the metal to bound electrons in materials like GaAs, GaP, or $LiNbO_3$ in turn also described as a set of nonlinear Lorentz oscillators. All electrons are assumed to be under the influence of electric and magnetic forces, so the dynamics that ensues in the metal and the dielectric contains surface and volume contributions simultaneously.

III. FREE ELECTRONS

An equation that describes free electrons inside the metal may be written as follows [3,40]:

$$m^*\frac{d\mathbf{v}}{dt}+\gamma m^*\mathbf{v}=e\mathbf{E}+\frac{e}{c}\mathbf{v}\times\mathbf{H}-\frac{\nabla p}{n}; \qquad (2)$$

where m^* is the effective mass of conduction electrons and n is their density; $\mathbf{v}$ is the electron velocity; $\mathbf{E}$ and $\mathbf{H}(=\mathbf{B})$ are electric and magnetic fields, respectively; and p is the electron

gas pressure. The full temporal derivative of the velocity in Eq. (2) can be written as

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}, \tag{3}$$

so Eq. (2) becomes

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} + \gamma\mathbf{v} = \frac{e}{m^*}\mathbf{E} + \frac{e}{m^* c}\mathbf{v}\times\mathbf{H} - \frac{\nabla p}{nm^*}. \tag{4}$$

Identifying the current density with $\mathbf{J} = ne\mathbf{v}$ makes it possible to rewrite Eq. (4) as

$$\frac{\partial \mathbf{J}}{\partial t} - \frac{\dot{n}}{n}\mathbf{J} + \mathbf{J}\cdot\nabla\left(\frac{\mathbf{J}}{ne}\right) + \gamma\mathbf{J} = \frac{ne^2}{m^*}\mathbf{E} + \frac{e}{m^* c}\mathbf{J}\times\mathbf{H} - \frac{\nabla p}{nm^*}. \tag{5}$$

After defining $\dot{\mathbf{P}}_j = \mathbf{J}$, Eq. (5) becomes

$$\ddot{\mathbf{P}}_j - \frac{\dot{n}}{n}\dot{\mathbf{P}}_j + (\dot{\mathbf{P}}_j\cdot\nabla)\left(\frac{\dot{\mathbf{P}}_j}{ne}\right) + \gamma\dot{\mathbf{P}}_j = \frac{ne^2}{m^*}\mathbf{E} + \frac{e}{m^* c}\dot{\mathbf{P}}_j\times\mathbf{H} - \frac{e\nabla p}{m^*}. \tag{6}$$

For free electrons the continuity equation $\dot{n}(\mathbf{r},t) = -\frac{1}{e}\nabla\cdot\dot{\mathbf{P}}_j$ supplements the equations of motion and may be integrated directly to yield

$$n(\mathbf{r},t) = n_0 - \frac{1}{e}\nabla\cdot\mathbf{P}_j, \tag{7}$$

where n_0 is the background, equilibrium charge density in the absence of any applied fields. In what follows our treatment departs from the usual procedure followed in the hydrodynamic model [3,19,40]. Assuming $\dot{n} \ll n$, the ratio $\dot{n}/n$ may be expanded in powers of $1/(n_0 e)$ to obtain

$$\frac{\dot{n}}{n} = -\frac{1}{n_0 e}\nabla\cdot\dot{\mathbf{P}}\left(1 - \frac{1}{n_0 e}\nabla\cdot\mathbf{P}\right)^{-1} \sim -\frac{\nabla\cdot\dot{\mathbf{P}}}{en_0} - \frac{1}{n_0^2 e^2}(\nabla\cdot\dot{\mathbf{P}})(\nabla\cdot\mathbf{P}) + \vartheta\left(\frac{1}{n_0^3 e^3}\right) + \cdots. \tag{8}$$

Substituting Eq. (8) back into Eq. (6) and neglecting terms of order $(1/n_0 e)^2$ and higher we get

$$\ddot{\mathbf{P}}_j + \gamma\dot{\mathbf{P}}_j = \frac{n_0 e^2}{m^*}\mathbf{E} - \frac{e}{m^*}\mathbf{E}(\nabla\cdot\mathbf{P}_j) + \frac{e}{m^* c}\dot{\mathbf{P}}_j\times\mathbf{H} - \frac{1}{n_0 e}[(\nabla\cdot\dot{\mathbf{P}}_j)\dot{\mathbf{P}}_j + (\dot{\mathbf{P}}_j\cdot\nabla)\dot{\mathbf{P}}_j] - \frac{e\nabla p}{m^*}. \tag{9}$$

In all the calculations that we performed for typical silver-based nanostructures ($n_0 \sim 5.8\times 10^{22}/\mathrm{cm}^3$) we consistently found $|\delta n| = |\nabla\cdot\mathbf{P}_j/e| \sim 10^{13}$–$10^{16}/\mathrm{cm}^3$. The lower bound is typical of uniform metal layers. The upper bound is characteristic of resonant subwavelength metallic slits and nanocavities [29].

The specific impact of pressure is seldom considered in the dynamics [40], but it is instructive to make a few observations in its regard. Pressure may be treated classically by assuming that electrons form an ideal gas, i.e., $p = nk_B T$ [52]. k_B is the Boltzmann constant and T is the temperature. The gradient of p in Eq. (9) becomes the gradient of n, which in turn may be related to the macroscopic polarization as follows:

$$-\frac{e\nabla p}{m^*} = -\frac{e}{m^*}k_B T\nabla\left(n_0 - \frac{1}{e}\nabla\cdot\mathbf{P}_j\right) = \frac{k_B T}{m^*}\nabla(\nabla\cdot\mathbf{P}_j). \tag{10}$$

It is interesting and equally instructive to also look at a quantum model of the pressure [53]. The quantum pressure is typically described as $p = p_0(n/n_0)^\gamma$, where $\gamma = (D+2)/D$ and D is the dimensionality of the problem [53,54]. For $D = 3$, we have $p = p_0(n/n_0)^{5/3}$, where $p_0 = n_0 E_F$, E_F is the Fermi energy, and n_0 is once again the equilibrium charge density. The leading pressure terms are

$$-\frac{e\nabla p}{m^*} = -\frac{ep_0}{m^* n_0^{5/3}}\frac{5}{3}n^{2/3}\nabla n = -\frac{5}{3}\frac{en_0 E_F}{m^* n_0^{5/3}}n^{2/3}\nabla n \approx \frac{5}{3}\frac{E_F}{m^*}\nabla(\nabla\cdot\mathbf{P}_j) - \frac{10}{9}\frac{E_F}{m^*}\frac{1}{n_0 e}(\nabla\cdot\mathbf{P}_j)\nabla(\nabla\cdot\mathbf{P}_j). \tag{11}$$

As already shown elsewhere using a two-fluid quantum model [53], and as Eq. (11) plainly suggests, the quantum model intrinsically contains a first-order classical, ideal electron gas contribution [if we equate the Fermi energy with $K_B T$ in Eq. (10)] and a nonlinear quantum correction of lower order. It is easy to see the impact of pressure if we scale the equations with respect to dimensionless time and longitudinal and transverse coordinates, $\tau = ct/\lambda_0$, $\xi = z/\lambda_0$, $\tilde{y} = y/\lambda_0$, respectively, where $\lambda_0 = 1\ \mu$m is arbitrarily chosen as the reference wavelength. As a result of this scaling Eq. (9) becomes

$$\ddot{\mathbf{P}}_j + \tilde{\gamma}\dot{\mathbf{P}}_j = \frac{n_0 e^2}{m^*}\left(\frac{\lambda_0}{c}\right)^2\mathbf{E} - \frac{e\lambda_0}{m^* c^2}\mathbf{E}(\nabla\cdot\mathbf{P}_j) + \frac{e\lambda_0}{m^* c^2}\dot{\mathbf{P}}_j\times\mathbf{H} - \frac{1}{n_0 e\lambda_0}[(\nabla\cdot\dot{\mathbf{P}}_j)\dot{\mathbf{P}}_j + (\dot{\mathbf{P}}_j\cdot\nabla)\dot{\mathbf{P}}_j] + \frac{5}{3}\frac{E_F}{m^* c^2}\nabla(\nabla\cdot\mathbf{P}_j) - \frac{10}{9}\frac{E_F}{m^* c^2}\frac{1}{n_0 e\lambda_0}(\nabla\cdot\mathbf{P}_j)\nabla(\nabla\cdot\mathbf{P}_j). \tag{12}$$

in addition to the magnetic Lorentz force, $(e\lambda_0/m^* c^2)\dot{\mathbf{P}}_j\times\mathbf{H}$, we have an explicit quadrupolelike [1] Coulomb term that arises from the continuity equation, $-(e\lambda_0/m^* c^2)\mathbf{E}(\nabla\cdot\mathbf{P}_j)$, convective terms proportional to $[(\nabla\cdot\dot{\mathbf{P}}_j)\dot{\mathbf{P}}_j + (\dot{\mathbf{P}}_j\cdot\nabla)\dot{\mathbf{P}}_j]$, and linear and nonlinear pressure terms proportional to $\nabla(\nabla\cdot\mathbf{P}_j)$ and $(\nabla\cdot\mathbf{P}_j)\nabla(\nabla\cdot\mathbf{P}_j)$, respectively. For silver, the Fermi velocity $v_F \sim 10^8$ cm/s so $(E_F/m^* c^2) \sim 10^{-5}$. This is then combined with $1/(n_0 e\lambda_0) \sim 10^{-10}$(cgs, silver). If, for the moment, we neglect all nonlinear contributions, Eq. (12) becomes

$$\ddot{\mathbf{P}}_j + \tilde{\gamma}\dot{\mathbf{P}}_j = \frac{n_0 e^2}{m^*}\left(\frac{\lambda_0}{c}\right)^2\mathbf{E} + \frac{5}{3}\frac{E_F}{m^* c^2}\nabla(\nabla\cdot\mathbf{P}_j). \tag{13}$$

Expanding the terms on the right hand side of Eq. (13) shows that the pressure couples orthogonal, free-electron polarization states and introduces a dynamical anisotropy. More generally, pressure could directly impact the linear dielectric function of the metal near its walls, should the fields become strongly

confined and/or their derivatives be large enough [55] (i.e., near sharp edges, corners, or in resonant subwavelength cavities) to introduce large, evanescent k vectors. The same is true for the nonlinear term: its magnitude could perturb Coulomb, Lorentz, or convective terms at high-enough intensity and/or if large-enough k vectors were excited. With these issues in mind, some simple considerations may be made about nonlinear sources derivable from Eq. (12). Assuming the incident pump is undepleted and time harmonic, lowest-order terms may be collected as follows:

$$\begin{aligned}\mathbf{P}^{\mathrm{SH}}_{\mathrm{NL,free}}(2\omega) \approx & -\frac{e\lambda_0}{m^*c^2}\mathbf{E}_\omega\nabla\cdot[\chi_{\mathrm{free}}(\omega)\mathbf{E}_\omega] \\ & -\frac{e\lambda_0}{m^*c^2}\chi_{\mathrm{free}}(\omega)i\beta\mathbf{E}_\omega\times\mathbf{H}_\omega \\ & -\frac{\beta^2\chi_{\mathrm{free}}(\omega)}{n_0e\lambda_0}\{\mathbf{E}_\omega\nabla\cdot[\chi_{\mathrm{free}}(\omega)\mathbf{E}_\omega] \\ & +(\mathbf{E}_\omega\cdot\nabla)\chi_{\mathrm{free}}(\omega)\mathbf{E}_\omega\}-\frac{10}{9}\frac{E_F}{m^*c^2}\frac{1}{n_0e\lambda_0} \\ & \times\{\nabla\cdot[\chi_{\mathrm{free}}(\omega)\mathbf{E}_\omega]\}\nabla\{\nabla\cdot[\chi_{\mathrm{free}}(\omega)\mathbf{E}_\omega]\},\end{aligned} \tag{14}$$

where $\beta = 2\pi\omega/\omega_0$ and $\chi_{\mathrm{free}}(\omega)$ is the free-electron portion of the dielectric function. A similar equation may be written for the third-harmonic polarization. Should the incident signal be a short pulse [25,30], material dispersion is included to all orders in Eq. (12) along with all appropriate boundary conditions [38,39]. The form of the spatial derivatives in Eq. (14) suggests that convective and quantum terms have properties similar to the Coulomb term. However, this could change for thin layers, subwavelength nanocavities, or for wavelengths in the near IR, visible, and UV ranges, where the fields can penetrate and become localized inside the metal [30,31]. For this reason our reference point for the dynamics of free electrons will be Eq. (12).

For localized ultrashort pulses harmonic generation occurs regardless of angle of incidence or polarization state. Additional dynamical features may be ascertained by decomposing Coulomb and Lorentz forces explicitly into all their harmonic components. The geometry of our system is shown in Fig. 2. Two fields of orthogonal polarizations and arbitrary amplitudes

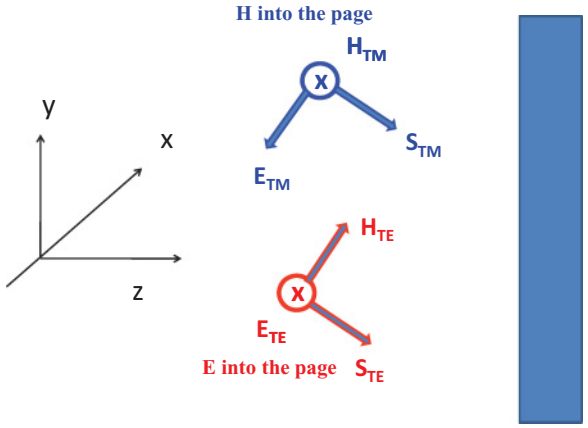

FIG. 2. (Color online) Incident TE- and TM-polarized fields. In all our cases the fields overlap but appear spatially separated for clarity.

are incident on a structure whose details are unspecified. The interaction takes place on the y-z plane and the fields are independent of x. Up to TH frequency, TE- and TM-polarized fields (according to Fig. 2) are as follows, respectively:

$$\begin{aligned}\mathbf{E}_{\mathrm{TE}} &= \mathbf{i}E_{\mathrm{TE}x} = \mathbf{i}\big[E^{\omega}_{\mathrm{TE}x}e^{-i\omega t}+\big(E^{\omega}_{\mathrm{TE}x}\big)^*e^{i\omega t}+E^{2\omega}_{\mathrm{TE}x}e^{-2i\omega t} \\ &\quad +\big(E^{2\omega}_{\mathrm{TE}x}\big)^*e^{2i\omega t}+E^{3\omega}_{\mathrm{TE}x}e^{-3i\omega t}+\big(E^{3\omega}_{\mathrm{TE}x}\big)^*e^{3i\omega t}\big]; \\ \mathbf{H}_{\mathrm{TE}} &= \begin{pmatrix}\mathbf{j}H_{\mathrm{TE}y}\\ +\mathbf{k}H_{\mathrm{TE}z}\end{pmatrix} = \mathbf{j}\big[H^{\omega}_{\mathrm{TE}y}e^{-i\omega t}+\big(H^{\omega}_{\mathrm{TE}y}\big)^*e^{i\omega t} \\ &\quad +H^{2\omega}_{\mathrm{TE}y}e^{-2i\omega t}+\big(H^{2\omega}_{\mathrm{TE}y}\big)^*e^{2i\omega t}+H^{3\omega}_{\mathrm{TE}y}e^{-3i\omega t} \\ &\quad +\big(H^{3\omega}_{\mathrm{TE}y}\big)^*e^{3i\omega t}\big]+\mathbf{k}\big[H^{\omega}_{\mathrm{TE}z}e^{-i\omega t}+\big(H^{\omega}_{\mathrm{TE}z}\big)^*e^{i\omega t} \\ &\quad +H^{2\omega}_{\mathrm{TE}z}e^{-2i\omega t}+\big(H^{2\omega}_{\mathrm{TE}z}\big)^*e^{2i\omega t}+H^{3\omega}_{\mathrm{TE}z}e^{-3i\omega t} \\ &\quad +\big(H^{3\omega}_{\mathrm{TE}z}\big)^*e^{3i\omega t}\big];\end{aligned} \tag{15}$$

$$\begin{aligned}\mathbf{E}_{\mathrm{TM}} &= \begin{pmatrix}\mathbf{j}E_{\mathrm{TM}y}\\ +\mathbf{k}E_{\mathrm{TM}z}\end{pmatrix} = \mathbf{j}\big[E^{\omega}_{\mathrm{TM}y}e^{-i\omega t}+\big(E^{\omega}_{\mathrm{TM}y}\big)^*e^{i\omega t} \\ &\quad +E^{2\omega}_{\mathrm{TM}y}e^{-2i\omega t}+\big(E^{2\omega}_{\mathrm{TM}y}\big)^*e^{2i\omega t}+E^{3\omega}_{\mathrm{TM}y}e^{-3i\omega t} \\ &\quad +\big(E^{3\omega}_{\mathrm{TM}y}\big)^*e^{3i\omega t}\big]+\mathbf{k}\big[E^{\omega}_{\mathrm{TM}z}e^{-i\omega t}+\big(E^{\omega}_{\mathrm{TM}z}\big)^*e^{i\omega t} \\ &\quad +E^{2\omega}_{\mathrm{TM}z}e^{-2i\omega t}+\big(E^{2\omega}_{\mathrm{TM}z}\big)^*e^{2i\omega t}+E^{3\omega}_{\mathrm{TM}z}e^{-3i\omega t} \\ &\quad +\big(E^{3\omega}_{\mathrm{TM}z}\big)^*e^{3i\omega t}\big]; \\ \mathbf{H}_{\mathrm{TM}} &= \mathbf{i}H_{\mathrm{TM}x} = \mathbf{i}\big[H^{\omega}_{\mathrm{TM}x}e^{-i\omega t}+\big(H^{\omega}_{\mathrm{TM}x}\big)^*e^{i\omega t}+H^{2\omega}_{\mathrm{TM}x}e^{-2i\omega t} \\ &\quad +\big(H^{2\omega}_{\mathrm{TM}x}\big)^*e^{2i\omega t}+H^{3\omega}_{\mathrm{TM}x}e^{-3i\omega t}+\big(H^{3\omega}_{\mathrm{TM}x}\big)^*e^{3i\omega t}\big].\end{aligned} \tag{16}$$

The field envelopes in Eqs. (15) and (16) are not assumed to be slowly varying and retain spatial and temporal dependence. The extraction of carrier frequencies is done only as a matter of convenience. With reference to Fig. 2, both TE- and TM-polarized fields have transverse components orthogonal to each other, but they should not be confused: one component points along $\hat{\mathbf{x}}$, the other along $\hat{\mathbf{y}}$. Equations (15) and (16) become

$$\mathbf{E} = \mathbf{i}E_{\mathrm{TE}x}+\mathbf{j}E_{\mathrm{TM}y}+\mathbf{k}E_{\mathrm{TM}z},\quad \mathbf{H} = \mathbf{i}H_{\mathrm{TM}x}+\mathbf{j}H_{\mathrm{TE}y}+\mathbf{k}H_{\mathrm{TE}z}. \tag{17}$$

In similar fashion the electric polarization has TE and TM components and may be expressed as

$$\mathbf{P} = (P_{\mathrm{TE}x}\mathbf{i}+P_{\mathrm{TM}y}\mathbf{j}+P_{\mathrm{TM}z}\mathbf{k}). \tag{18}$$

An expansion of the magnetic Lorentz force using Eqs. (17) and (18) yields the following constituents:

$$\begin{aligned}\dot{\mathbf{P}}\times\mathbf{H} &= (\dot{P}_{\mathrm{TE}x}\mathbf{i}+\dot{P}_{\mathrm{TM}y}\mathbf{j}+\dot{P}_{\mathrm{TM}z}\mathbf{k}) \\ &\quad\times(H_{\mathrm{TM}x}\mathbf{i}+H_{\mathrm{TE}y}\mathbf{j}+H_{\mathrm{TE}z}\mathbf{k}) \\ &= (\dot{P}_{\mathrm{TM}y}H_{\mathrm{TE}z}-\dot{P}_{\mathrm{TM}z}H_{\mathrm{TE}y})\mathbf{i} \\ &\quad+(\dot{P}_{\mathrm{TM}z}H_{\mathrm{TM}x}-\dot{P}_{\mathrm{TE}x}H_{\mathrm{TE}z})\mathbf{j} \\ &\quad+(\dot{P}_{\mathrm{TE}x}H_{\mathrm{TE}y}-\dot{P}_{\mathrm{TM}y}H_{\mathrm{TM}x})\mathbf{k}.\end{aligned} \tag{19}$$

These equations show that even a normally incident, TE-polarized field generates a nonzero, TM-polarized harmonic component via the term $(\dot{P}_{\mathrm{TE}x}H_{\mathrm{TE}y})\mathbf{k}$, thus opening a possible catalytic interaction channel. At oblique incidence the term $(-\dot{P}_{\mathrm{TE}x}H_{\mathrm{TE}z})\mathbf{j}$ also provides nonzero gain for a TM-polarized

harmonic signal. The Coulomb term may also be decomposed as follows:

$$(\nabla \cdot \mathbf{P}_j)\mathbf{E} = \left(\frac{\partial P_{\mathrm{TM}z}}{\partial z} + \frac{\partial P_{\mathrm{TM}y}}{\partial y}\right)(\mathbf{i}E_{\mathrm{TE}x} + \mathbf{j}E_{\mathrm{TM}y} + \mathbf{k}E_{\mathrm{TM}z}), \tag{20}$$

where $\partial P_{\mathrm{TE}x}/\partial x = 0$. Therefore an incident beam with a mixed polarization state activates all possible conversion channels. Simultaneously TE- and TM-polarized incident fields lead to efficiencies that may differ and be much enhanced compared to pumping with either TE- or TM-polarized light only.

IV. BOUND ELECTRONS

Bound electrons differ from free electrons in at least two ways: (i) they may be under the action of linear and nonlinear restoring forces and (ii) the average local charge density remains constant in time, as electrons are not free to leave their atomic sites. The method that we present here to treat Coulomb and Lorentz forces was derived previously for SHG [4] and plane waves in the undepleted pump regime. Here we generalize that approach to treat pulses of arbitrary duration up to the TH frequency. The method may be combined with nonlinear restoring forces ($\chi^{(2)}$, $\chi^{(3)}$, etc.) to generalize the nonlinear response of dielectrics or semiconductors to include surface phenomena dynamically. The ability to describe the simultaneous excitation of linear and nonlinear plasmonic phenomena is important to model nonlinear semiconductor nanocavities or slits in the UV range, where the dielectric function may be negative [56].

Neglecting for the moment nonlinear restoring forces, Newton's second law for one species of core electrons leads to an effective polarization equation for bound charges that reads as follows:

$$\ddot{\mathbf{P}}_b + \gamma_b \dot{\mathbf{P}}_b + \omega_{0,b}^2 \mathbf{P}_b = \frac{n_{0,b}e^2}{m_b^*}\mathbf{E} + \frac{e}{m_b^* c}\dot{\mathbf{P}}_b \times \mathbf{H}, \tag{21}$$

where $\mathbf{P}_b = n_{0,b}e\mathbf{r}_b$ is the polarization, $\mathbf{r}_b$ is the electron's position relative to an equilibrium origin, $n_{0,b}$ is the constant density, m_b^* is the bound electron's effective mass, and $\dot{\mathbf{P}}_b = n_{0,b}e\dot{\mathbf{r}}_b$ is the bound current density. Up to the third-harmonic frequency, the fields may be written as

$$\begin{aligned}\mathbf{E} = (&\mathbf{E}_\omega e^{i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} + \mathbf{E}_\omega^* e^{-i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} + \mathbf{E}_{2\omega} e^{2i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} \\ &+\mathbf{E}_{2\omega}^* e^{-2i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} + \mathbf{E}_{3\omega} e^{3i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} + \mathbf{E}_{3\omega}^* e^{-3i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)}), \\ \mathbf{H} = (&\mathbf{H}_\omega e^{i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} + \mathbf{H}_\omega^* e^{-i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} + \mathbf{H}_{2\omega} e^{2i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} \\ &+\mathbf{H}_{2\omega}^* e^{-2i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} + \mathbf{H}_{3\omega} e^{3i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)} + \mathbf{H}_{3\omega}^* e^{-3i(\mathbf{k}\cdot\mathbf{r}_b-\omega t)}).\end{aligned} \tag{22}$$

Expanding the fields in powers of $\mathbf{k}\cdot\mathbf{r}_b$ we have [4]

$$\mathbf{E} = \begin{pmatrix} \mathbf{E}_\omega e^{-i\omega t}\left(1 + i\mathbf{k}\cdot\mathbf{r}_b + \frac{(i\mathbf{k}\cdot\mathbf{r}_b)^2}{2} + \cdots\right) + \mathbf{E}_\omega^* e^{i\omega t}\left(1 - i\mathbf{k}\cdot\mathbf{r}_b + \frac{(-i\mathbf{k}\cdot\mathbf{r}_b)^2}{2} + \cdots\right) \\ +\mathbf{E}_{2\omega} e^{-2i\omega t}\left(1 + 2i\mathbf{k}\cdot\mathbf{r}_b + \frac{(2i\mathbf{k}\cdot\mathbf{r}_b)^2}{2} + \cdots\right) + \mathbf{E}_{2\omega}^* e^{2i\omega t}\left(1 - 2i\mathbf{k}\cdot\mathbf{r}_b + \frac{(-2i\mathbf{k}\cdot\mathbf{r}_b)^2}{2} + \cdots\right) \\ +\mathbf{E}_{3\omega} e^{-3i\omega t}\left(1 + 3i\mathbf{k}\cdot\mathbf{r}_b + \frac{(3i\mathbf{k}\cdot\mathbf{r}_b)^2}{2} + \cdots\right) + \mathbf{E}_{3\omega}^* e^{3i\omega t}\left(1 - 3i\mathbf{k}\cdot\mathbf{r}_b + \frac{(-3i\mathbf{k}\cdot\mathbf{r}_b)^2}{2} + \cdots\right) \end{pmatrix}, \tag{23}$$

and similarly for the magnetic field. The solutions for the electron's position and its derivatives are

$$\mathbf{r}_b = \mathbf{r}_\omega e^{-i\omega t} + \mathbf{r}_{2\omega} e^{-2i\omega t} + \mathbf{r}_{3\omega} e^{-3i\omega t} + \text{c.c.}, \tag{24}$$

$$\begin{aligned}\dot{\mathbf{r}}_b = &(\dot{\mathbf{r}}_\omega - i\omega\mathbf{r}_\omega)e^{-i\omega t} + (\dot{\mathbf{r}}_{2\omega} - 2i\omega\mathbf{r}_{2\omega})e^{-2i\omega t} \\ &+ (\dot{\mathbf{r}}_{3\omega} - 3i\omega\mathbf{r}_{3\omega})e^{-3i\omega t} + \text{c.c.},\end{aligned} \tag{25}$$

$$\begin{aligned}\ddot{\mathbf{r}}_b = &(\ddot{\mathbf{r}}_\omega - 2i\omega\dot{\mathbf{r}}_\omega - \omega^2\mathbf{r}_\omega)e^{-i\omega t} \\ &+ (\ddot{\mathbf{r}}_{2\omega} - 4i\omega\dot{\mathbf{r}}_{2\omega} - 4\omega^2\mathbf{r}_{2\omega})e^{-2i\omega t} \\ &+ (\ddot{\mathbf{r}}_{3\omega} - 6i\omega\dot{\mathbf{r}}_\omega - 9\omega^2\mathbf{r}_\omega)e^{-3i\omega t} + \text{c.c.}\end{aligned} \tag{26}$$

As before, the envelope functions defined in Eqs. (22)–(26) are not assumed to be slowly varying as all temporal and spatial derivatives will be retained. Substitution of Eqs. (22)–(26) into Eq. (21) leads to

$$\begin{aligned}\ddot{\mathbf{r}}_\omega + (\gamma_b - 2i\omega)\dot{\mathbf{r}}_\omega + \left(\omega_0^2 - \omega^2 + i\gamma_b\omega\right)\mathbf{r}_\omega = &\frac{e}{m_b^*}\begin{pmatrix} \mathbf{E}_\omega \\ -i\mathbf{k}\cdot\mathbf{r}_{2\omega}\mathbf{E}_\omega^* \\ -2i\mathbf{k}\cdot\mathbf{r}_{3\omega}\mathbf{E}_{2\omega}^* \\ +2i\mathbf{k}\cdot\mathbf{r}_\omega^*\mathbf{E}_{2\omega} \\ +3i\mathbf{k}\cdot\mathbf{r}_{2\omega}^*\mathbf{E}_{3\omega} \end{pmatrix} \\ &+ \frac{e}{m_b^* c}\begin{pmatrix} (\dot{\mathbf{r}}_\omega^* + i\omega\mathbf{r}_\omega^*) \times \mathbf{H}_{2\omega} + (\dot{\mathbf{r}}_{2\omega} - 2i\omega\mathbf{r}_{2\omega}) \times \mathbf{H}_\omega^* \\ (\dot{\mathbf{r}}_{2\omega}^* + 2i\omega\mathbf{r}_{2\omega}^*) \times \mathbf{H}_{3\omega} + (\dot{\mathbf{r}}_{3\omega} - 3i\omega\mathbf{r}_{3\omega}) \times \mathbf{H}_{2\omega}^* \\ -(i\dot{\mathbf{r}}_\omega + \omega\mathbf{r}_\omega) \times \mathbf{H}_\omega^*\mathbf{k}\cdot\mathbf{r}_\omega - (-i\dot{\mathbf{r}}_\omega^* + \omega\mathbf{r}_\omega^*) \times \mathbf{H}_\omega\mathbf{k}\cdot\mathbf{r}_\omega \\ +(i\dot{\mathbf{r}}_\omega + \omega\mathbf{r}_\omega) \times \mathbf{H}_\omega\mathbf{k}\cdot\mathbf{r}_\omega^* \end{pmatrix},\end{aligned}$$

$$\ddot{\mathbf{r}}_{2\omega}+(\gamma_b-4i\omega)\dot{\mathbf{r}}_{2\omega}+\left(\omega_0^2-4\omega^2+2i\gamma_b\omega\right)\mathbf{r}_{2\omega}=\frac{e}{m_b^*}\begin{pmatrix}+\mathbf{E}_{2\omega}\\ +i\mathbf{k}\cdot\mathbf{r}_\omega\mathbf{E}_\omega\\ -i\mathbf{k}\cdot\mathbf{r}_{3\omega}\mathbf{E}_\omega^*\\ +3i\mathbf{k}\cdot\mathbf{r}_\omega^*\mathbf{E}_{3\omega}\end{pmatrix}+\frac{e}{m_b^*c}\begin{pmatrix}(\dot{\mathbf{r}}_\omega-i\omega\mathbf{r}_\omega)\times\mathbf{H}_\omega\\ +(\dot{\mathbf{r}}_{3\omega}-3i\omega\mathbf{r}_{3\omega})\times\mathbf{H}_\omega^*+(\dot{\mathbf{r}}_\omega^*+i\omega\mathbf{r}_\omega^*)\times\mathbf{H}_{3\omega}\\ -(i\dot{\mathbf{r}}_{2\omega}+2\omega\mathbf{r}_{2\omega})\times\mathbf{H}_\omega^*\mathbf{k}\cdot\mathbf{r}_\omega-2(i\dot{\mathbf{r}}_\omega+\omega\mathbf{r}_\omega)\times\mathbf{H}_\omega^*\mathbf{k}\cdot\mathbf{r}_{2\omega}\\ -(-i\dot{\mathbf{r}}_\omega^*+\omega\mathbf{r}_\omega^*)\times\mathbf{H}_\omega\mathbf{k}\cdot\mathbf{r}_{2\omega}-2(-i\dot{\mathbf{r}}_\omega^*+\omega\mathbf{r}_\omega^*)\times\mathbf{H}_{2\omega}\mathbf{k}\cdot\mathbf{r}_\omega\\ +(i\dot{\mathbf{r}}_{2\omega}+2\omega\mathbf{r}_{2\omega})\times\mathbf{H}_\omega\mathbf{k}\cdot\mathbf{r}_\omega^*+2(i\dot{\mathbf{r}}_\omega+\omega\mathbf{r}_\omega)\times\mathbf{H}_{2\omega}\mathbf{k}\cdot\mathbf{r}_\omega^*\end{pmatrix},$$

$$\ddot{\mathbf{r}}_{3\omega}+(\gamma_b-6i\omega)\dot{\mathbf{r}}_{3\omega}+\left(\omega_0^2-9\omega^2+3i\gamma_b\omega\right)\mathbf{r}_{3\omega}=\frac{e}{m_b^*}\begin{pmatrix}+\mathbf{E}_{3\omega}\\ +i\mathbf{k}\cdot\mathbf{r}_{2\omega}\mathbf{E}_\omega\\ +2i\mathbf{k}\cdot\mathbf{r}_\omega\mathbf{E}_{2\omega}\end{pmatrix}+\frac{e}{m_b^*c}\begin{pmatrix}(\dot{\mathbf{r}}_{2\omega}-2i\omega\mathbf{r}_{2\omega})\times\mathbf{H}_\omega+(\dot{\mathbf{r}}_\omega-i\omega\mathbf{r}_\omega)\times\mathbf{H}_{2\omega}\\ +(i\dot{\mathbf{r}}_\omega+\omega\mathbf{r}_\omega)\times\mathbf{H}_\omega\mathbf{k}\cdot\mathbf{r}_\omega\end{pmatrix}. \tag{27}$$

In writing Eqs. (27) we have excluded higher-order terms that couple magnetic and electric dipoles, terms like $[(i\dot{\mathbf{r}}_\omega+\omega\mathbf{r}_\omega)\times\mathbf{H}_\omega\mathbf{k}\cdot\mathbf{r}_\omega]$ that already appear in Eq. (27) but contain at least two harmonic fields. These terms could become important in solids at TW/cm^2 or highly nonlinear plasmas [57,58] or for extremely short, intense pulses. Of course, one should consider these and other terms if different pumping conditions are used, and matters should be evaluated on a case-by-case basis [59]. Finally, Eq. (27) may be simplified if we neglect higher-order magnetic dipole-electric dipole terms, and if we identify

$$\begin{gathered}\mathbf{P}_{b,\omega}=n_{0,b}e\mathbf{r}_\omega,\quad \mathbf{P}_{b,2\omega}=n_{0,b}e\mathbf{r}_{2\omega},\quad \mathbf{P}_{b,3\omega}=n_{0,b}e\mathbf{r}_{3\omega},\\ i\mathbf{k}\cdot n_{0,b}e\mathbf{r}_\omega\approx\nabla\cdot\mathbf{P}_{b,\omega},\quad 2i\mathbf{k}\cdot n_{0,b}e\mathbf{r}_{2\omega}\approx\nabla\cdot\mathbf{P}_{b,2\omega},\quad 3i\mathbf{k}\cdot n_{0,b}e\mathbf{r}_{3\omega}\approx\nabla\cdot\mathbf{P}_{b,3\omega}.\end{gathered} \tag{28}$$

Then, Eq. (21) becomes

$$\ddot{\mathbf{P}}_{b,\omega}+\tilde{\gamma}_{b,\omega}\dot{\mathbf{P}}_{b,\omega}+\tilde{\omega}_{0,b,\omega}^2\mathbf{P}_{b,\omega}\approx\frac{n_{0,b}e^2\lambda_0^2}{m_b^*c^2}\mathbf{E}_\omega+\frac{e\lambda_0}{m_b^*c^2}\begin{pmatrix}-\frac{1}{2}\mathbf{E}_\omega^*\nabla\cdot\mathbf{P}_{b,2\omega}\\ +2\mathbf{E}_{2\omega}\nabla\cdot\mathbf{P}_{b,\omega}^*\\ -\frac{2}{3}\mathbf{E}_{2\omega}^*\nabla\cdot\mathbf{P}_{b,3\omega}\\ -\frac{3}{2}\mathbf{E}_{3\omega}\nabla\cdot\mathbf{P}_{b,2\omega}^*\end{pmatrix}+\frac{e\lambda_0}{m_b^*c^2}\begin{pmatrix}(\dot{\mathbf{P}}_{b,\omega}^*+i\omega\mathbf{P}_{b,\omega}^*)\times\mathbf{H}_{2\omega}\\ +(\dot{\mathbf{P}}_{b,2\omega}-2i\omega\mathbf{P}_{b,2\omega})\times\mathbf{H}_\omega^*\\ +(\dot{\mathbf{P}}_{b,2\omega}^*+2i\omega\mathbf{P}_{b,2\omega}^*)\times\mathbf{H}_{3\omega}\\ +(\dot{\mathbf{P}}_{b,3\omega}-3i\omega\mathbf{P}_{b,3\omega})\times\mathbf{H}_{2\omega}^*\end{pmatrix},$$

$$\ddot{\mathbf{P}}_{b,2\omega}+\tilde{\gamma}_{b,2\omega}\dot{\mathbf{P}}_{b,2\omega}+\tilde{\omega}_{0,b,2\omega}^2\mathbf{P}_{b,2\omega}\approx\frac{n_{0,b}e^2\lambda_0^2}{m_b^*c^2}\mathbf{E}_{2\omega}+\frac{e\lambda_0}{m_b^*c^2}\begin{pmatrix}\mathbf{E}_\omega\nabla\cdot\mathbf{P}_{b,\omega}\\ -\frac{1}{3}\mathbf{E}_\omega^*\nabla\cdot\mathbf{P}_{b,3\omega}\\ -3\mathbf{E}_{3\omega}\nabla\cdot\mathbf{P}_{b,\omega}^*\end{pmatrix}+\frac{e\lambda_0}{m_b^*c^2}\begin{pmatrix}(\dot{\mathbf{P}}_{b,\omega}-i\omega\mathbf{P}_{b,\omega})\times\mathbf{H}_\omega\\ +(\dot{\mathbf{P}}_{b,\omega}^*+i\omega\mathbf{P}_{b,\omega}^*)\times\mathbf{H}_{3\omega}\\ +(\dot{\mathbf{P}}_{b,3\omega}-3i\omega\mathbf{P}_{b,3\omega})\times\mathbf{H}_\omega^*\end{pmatrix},$$

$$\ddot{\mathbf{P}}_{b,3\omega}+\tilde{\gamma}_{b,3\omega}\dot{\mathbf{P}}_{b,3\omega}+\tilde{\omega}_{0,b,3\omega}^2\mathbf{P}_{b,3\omega}\approx\frac{n_{0,b}e^2\lambda_0^2}{m_b^*c^2}\mathbf{E}_{3\omega}+\frac{e\lambda_0}{m_b^*c^2}\begin{pmatrix}\frac{1}{2}\mathbf{E}_\omega\nabla\cdot\mathbf{P}_{b,2\omega}\\ +2\mathbf{E}_{2\omega}\nabla\cdot\mathbf{P}_{b,\omega}\end{pmatrix}+\frac{e\lambda_0}{m_b^*c^2}\begin{pmatrix}(\dot{\mathbf{P}}_{b,2\omega}-2i\omega\mathbf{P}_{b,2\omega})\times\mathbf{H}_\omega\\ +(\dot{\mathbf{P}}_{b,\omega}-i\omega\mathbf{P}_{b,\omega})\times\mathbf{H}_{2\omega}\end{pmatrix}. \tag{29}$$

The scaled coefficients are $\tilde{\gamma}_{b,N\omega}=(\gamma_b-Ni\omega)$, $\tilde{\omega}_{0,b,N\omega}^2=(\omega_0^2-(N\omega)^2+i\gamma_bN\omega)$, where N is an integer that denotes the given harmonic. We emphasize that all envelope functions in Eq. (29) are generic and allowed to vary rapidly in space and time, as demonstrated by the presence of spatial and temporal derivatives up to all orders of dispersion. Expanding Eq. (12) into all its harmonics with envelope functions and oscillating factors leads to equations that are very similar to Eqs. (29). Equations (12) and (29) thus form a set of coupled equations that describe free and bound charges that give rise to second- and third-harmonic generation in metallic structures of arbitrary geometry and are valid in the ultrashort pulse regime. Equation (29) is applicable to dielectrics and semiconductors alike, for depleted pumps and

harmonic down-conversion or other cases where the harmonic fields may be more intense than the fundamental field. The introduction of nonlinear restoring forces generalizes Eq. (29) to describe nonlinear phenomena in dielectric and semiconductor nanocavities. This may be achieved by adding nonlinear terms to the right-hand sides of each of Eq. (29) or by introducing a nonlinear polarization in the usual way:

$$\mathbf{P}_{\mathrm{NL}} = \chi^{(2)}\mathbf{E}^2 + \chi^{(3)}\mathbf{E}^3 + \cdots. \tag{30}$$

A comparison between Eqs. (12) and (29) shows that at this order of approximation the equations have similar form, even though the Coulomb terms have different origins. For free electrons the equation of continuity accounts for charge density variations in time. In the case of bound charges it is the spatial variation of the fields that leads to similar contributions. The Coulomb terms have opposite signs, negative for free electrons and positive for bound charges, with identifiable, effective charge distributions and currents that add to and interfere with their free-electron counterparts. Given the similarities between Eqs. (12) and (29), it is reasonable to expect comparable qualitative responses from free and bound charges [7,11]. We also note that it is possible to arrive at slightly different forms of Eqs. (29) that involve the fields explicitly rather than the polarization, as was done in Ref. [4].

We conclude this section with a few words about expectations from the model. Equations (12) and (29) are solved simultaneously with Maxwell's equation for pulses of arbitrary duration without approximations. The coupled equations that result from expanding up to the TH frequency (92 in all) are integrated using a fast Fourier transform, time domain, split-step beam propagation method. The technique is outlined in Ref. [60], where it was originally developed for the wave equation to describe slowly varying envelope functions in order to integrate an equation of motion that was first order in time. However, since Maxwell's equations are first order in time at the start, the method is easily extended by removing all approximations to take into account all orders of reflection (spatial derivatives) and dispersion (temporal derivatives). The integration scheme is stable, electric and magnetic fields and their derivatives are colocated, and the evaluation of spatial derivatives across hard interfaces is straightforward. Integration times range from a few minutes to a few hours on a 2.4-GHz Pentium Core 2 Duo, depending on pulse duration and angle of approach.

As may be easily ascertained from Eqs. (12) and (29), the free parameters the model allows are the effective free and bound electron masses and densities. These factors are important because they appear as a ratio (plasma frequency) and also singly in the nonlinear terms to completely determine the quantitative aspect of conversion efficiencies. Although the values of electron mass and density are well known, in practice strain during the deposition process causes these values to fluctuate and so the issue has been addressed often in the literature [61–64]. The actual values not only appear to be sensitive to the method of growth and layer thickness [64] but also depend on the natural porosity of the metal under consideration and the surface on which it is deposited [65,66]. Reference [64] reports that using dc magnetron sputtering, layers of gold exhibited effective conduction electron masses that ranged from $0.48m_e$ to $1.14m_e$ for 9.6- and 22-nm layers, respectively. Similar results were found for silver, with deduced effective masses that varied from $1.06m_e$ to $1.99m_e$ for 10- and 19-nm layers, respectively. The situation improves for rf sputtering, but effective masses still vary with layer thickness between 20% for silver (from $1.02m_e$ to $0.84m_e$) and 35% for gold (from $1.38m_e$ to $1.04m_e$). As a consequence of these relatively large fluctuations, one should expect that electron density, plasma frequency, skin depth, and Fermi energy will also vary accordingly, allowing substantial flexibility in departing from tabulated values.

Although porosity can be used to control plasmonic behavior of gold across a large wavelength range [66], it can also manifest itself in a somewhat disquieting manner [65]. The experimental results reported in Ref. [65] indicated that a silver layer deposited on a glass substrate may be much more porous on the air side than on the glass side. By monitoring surface plasmons propagating on either side, it was determined that for an incident wavelength of 632 nm the real part of the dielectric "constant" on the air side was approximately 25% smaller compared to the glass side, while the imaginary part could vary by as much as 60%! These differences are significant, persist for thick layers, and ultimately modify the quantitative aspects of linear propagation (actual reflection and transmission) and harmonic generation. Finally, there is the issue of surface roughness. It is easy to imagine how in the linear regime roughness could act in ways similar to porosity, in that it could lead to changes of the effective mass, density, skin depth, and Fermi energy. In the nonlinear regime the enhancement of SHG due to surface roughness has already been discussed [67]: it takes place in a manner similar to what occurs in the case of surface enhanced Raman scattering [55], and as such it is also characterized by the formation of SH hot spots. Although under normal conditions all these issues may be neglected in favor of qualitative rather than quantitative agreement (e.g., uniform layers), the discussion above provides hints that sometimes the interpretations of linear and/or nonlinear plasmonic phenomena that revolve around subwavelength slits, holes, sharp edges, or extremely narrow, nanometer-wide plasmonic transmission resonances or band gaps may be more subtle and intricate than previously thought.

V. THE IMPACT OF CONVECTION

We now turn to the assessment of convective terms on harmonic generation. Figure 3 shows the calculated SHG efficiency on reflection from a 125-nm-thick silver film suspended in air of a TM-polarized, $\sim$1-GW/cm^2 pump pulse approximately 50 fs in duration, tuned to 1064 nm. We measure efficiency as the ratio of (total energy converted) to (total incident pump energy) after the pulse has passed and show results with and without convective contributions. The detected radiation is also TM polarized. In the case of a single, uniform thick layer (see caption of Fig. 3 for the parameters used in the calculation) the quantitative effect of convection at 1064 nm is evidently to increase conversion efficiency by a factor of $\sim$2.5. We note that conversion efficiency quickly converges for pulses just a few optical cycles in duration, so the same results are obtained for longer pulses. Even though in Fig. 3 conversion efficiency is enhanced, convective

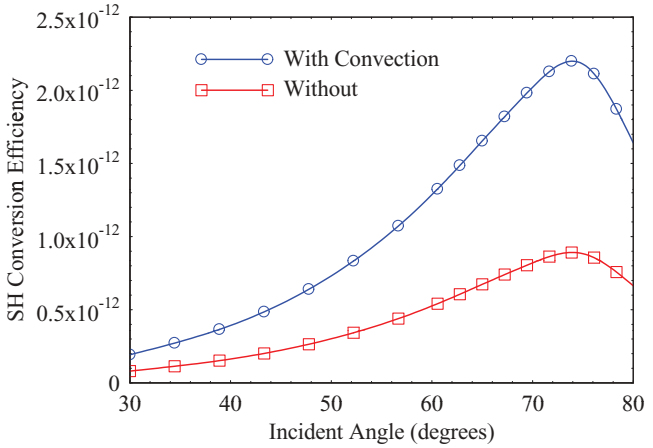

FIG. 3. (Color online) Conversion efficiency for TM-incident (1064 nm)–TM-detected (532 nm) light for a 125-nm-thick silver layer. The effective mass and densities of both free and bound charges are taken to be $m^* = m_e$ and $n_0 = 5.8 \times 10^{22}/\mathrm{cm}^3$, respectively.

contributions are sensitive to the geometry of the structure and their impact cannot be generalized as simply as it might appear from Fig. 3. As an example we calculated transmitted and reflected second-harmonic conversion efficiencies from a transparent metal/dielectric multilayer stack similar to those studied in Refs. [30] and [31]. The structure consists of a five-period Ta_2O_5 (125-nm-thick)/Ag (20-nm-thick) stack, with an incident, $\sim$1 GW/cm^2, 60-fs pulse tuned to 800 nm. The stack and its plane-wave transmission at normal incidence are depicted in Fig. 4. In Fig. 5 we report the predicted transmitted and reflected conversion efficiencies for TM-incident or TM-detected second-harmonic generation vs incident angle, with and without convective terms. The model we use presently is much improved compared to the model that was employed in Ref. [30], where nonlinear sources were derived only from the Lorentz force. Also significant is the fact that in this model dielectric interfaces become active

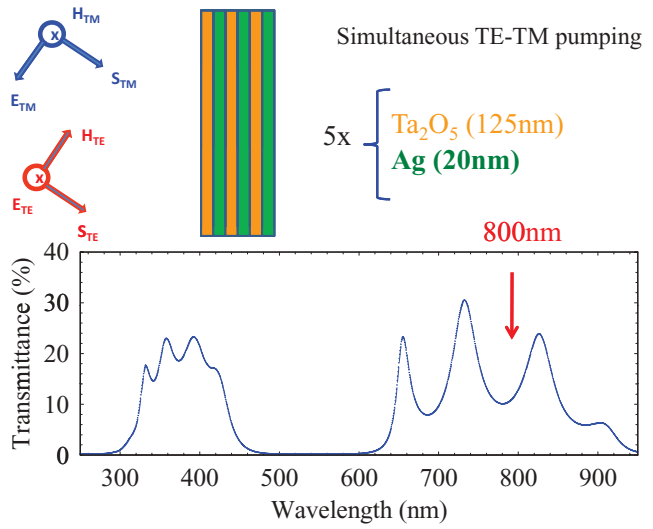

FIG. 4. (Color online) (Bottom) Transmission at normal incidence from a five-period Ta_2O_5/Ag multilayer stack, having layer thicknesses of 125 and 20 nm, respectively. (Top) Picture of the incident pulse and the stack.

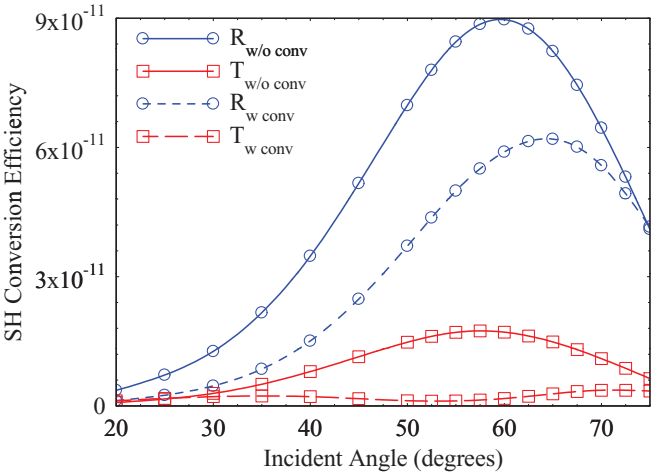

FIG. 5. (Color online) TM-incident–TM-detected, reflected, and transmitted SH efficiencies, with and without convection. Convection shifts and reduces the amplitude of the reflected signal, while turning the transmission maximum into a minimum. The subscript w/o stands for "without," while the subscript w stands for "with." We have assumed that $m^* = m_e$ and $n_0 = 5.8 \times 10^{22}/\mathrm{cm}^3$.

via the presence of the spatial derivatives of the polarization in Eq. (29), that explicitly account for symmetry breaking contributions from all bound electron sources. In other words, this stack would still generate harmonics via Eq. (29) even if we turned off all the nonlinear terms in Eq. (12). Our calculations thus suggest that in this case convection tends to reduce and shift the peaks of conversion efficiency. With convection the reflected SH signal is smaller compared to its counterpart without convection by $\sim$50% and the peak is shifted toward larger angles. For the transmitted SH signal the tendency is more dramatic compared to the single layer: the maximum turns into a minimum. Although these trends could change for different layer thicknesses, this example should suffice to highlight the sensitivity of the process to convective terms. The quantitative consequences of convection are far more dramatic if the stack of Fig. 4 is pumped with a mixed TE–TM polarization state having equal TE- and TM-polarized incident field amplitudes and TE-polarized detected signal. The results are depicted in Fig. 6. The figure shows that in addition to shifting the peaks to larger angles, the inclusion of convective terms can also reduce reflected and transmitted second-harmonic conversion efficiencies in this case by a factor of $\sim$50. Usually neither a single TE- nor a single TM-polarized incident field can generate a TE-polarized SH signal without additional degrees of freedom and/or nonlinear sources. However, the dual-pumping mode opens up and couples all available interaction channels, so TE- and TM-polarized SH and TH light may be efficiently produced. These results show that in the general case the presence of convection can influence the dynamics and strongly impact the results quantitatively and qualitatively. Field penetration and localization inside the metal may accentuate its importance.

VI. EFFECTS DUE TO BOUND CHARGES

We now change the context and explore the influence of bound charges. In Fig. 7 we show SH (400 nm) and

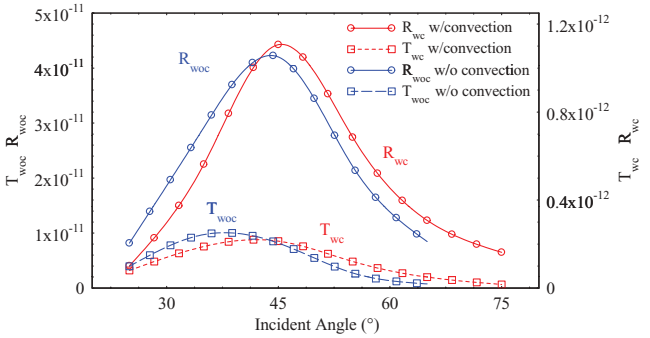

FIG. 6. (Color online) TE- or TM-incident and TE-detected, reflected, and transmitted SH conversion efficiencies with (right axis) and without (left axis) convection. Convection shifts the peaks to larger angles but reduces both signals by nearly a factor of 50. The subscript woc stands for "without convection" and wc stands for "with convection." $m^* = m_e$ and $n_0 = 5.8 \times 10^{22}/\text{cm}^3$.

TH (266 nm) emission patterns that originate from a silver nanopillar having a square cross section 200 nm on each side. For simplicity we have assumed that in the metal $\chi^{(3)} = 0$. A TM-polarized, 20-fs pump pulse is incident from the left. The detected harmonic fields are also TM polarized. In Figs. 7(a) and 7(b) we show snapshots of SH and TH emission patterns that emerge from the sample if only free charge contributions were turned on [Drude term only in Eq. (1)]. In Figs. 7(c) and 7(d) we show the corresponding field patterns when both free and bound charges contribute simultaneously. We can see that bound charges can impact field patterns, intensities, and conversion efficiencies. We now recalculate harmonic field emission patterns for a TE-polarized incident pulse, otherwise similar in nature to the pulse used in Fig. 7, and plot the results in Fig. 8. Again we find that the simultaneous presence of free and bound charges can significantly alter the total SH and TH conversion efficiencies and respective field patterns. In all cases adjusting the relative weights of free to bound electron masses and densities will also modify the relative contributions and affect the results. Although it may be possible to adjust effective surface and volume coefficients to the extent of obtaining a reasonable fit with experimental results, our simulations suggest that bound charges interfere with free charges and that ignoring their contributions may amount to an important omission in the larger, more complicated dynamical context outlined above.

VII. ELECTRON GAS PRESSURE

From Eqs. (12) and (13) above one may determine that the potential for dynamical modifications of the linear dielectric

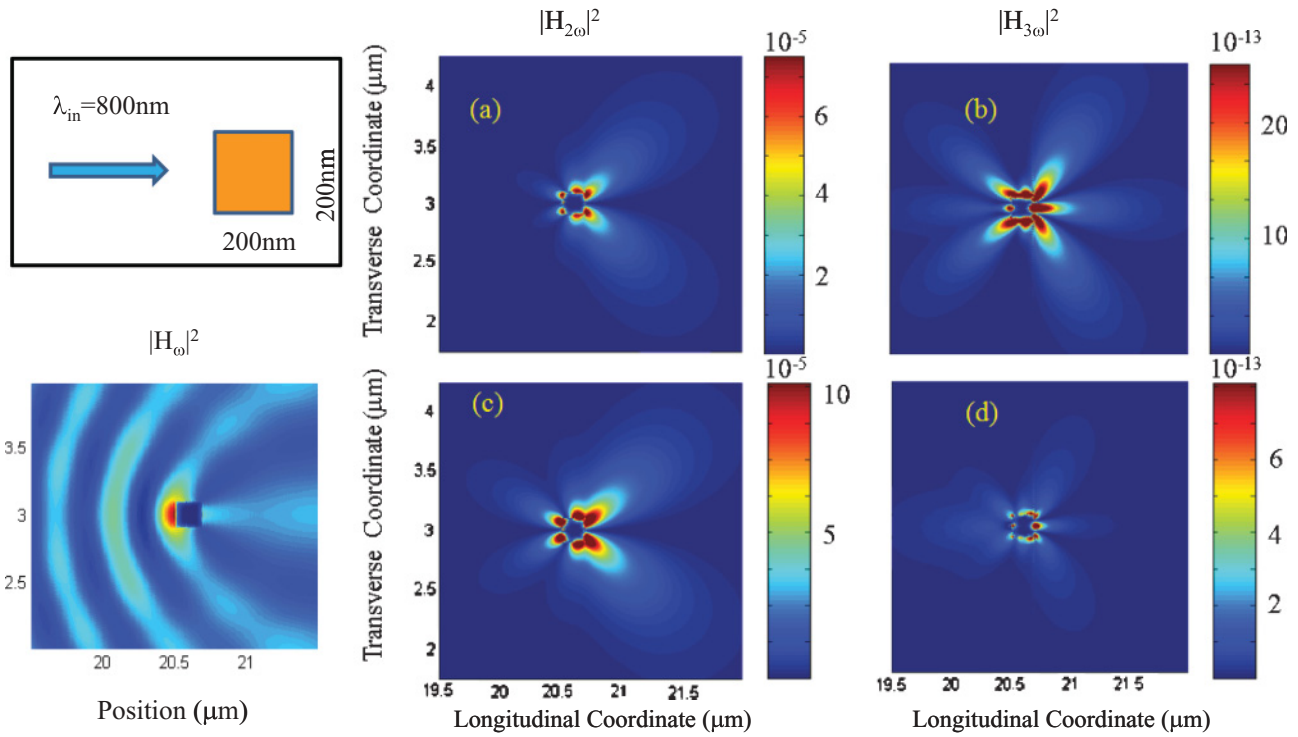

FIG. 7. (Color online) (Right) A TM-polarized 20-fs pump pulse tuned to 800 nm is incident from the left on a silver nanopillar 200 nm on each side. All detected fields are TM polarized. The snapshots are recorded when the peak of the pump pulse reaches the object. The spatial extension of the pulse is hundreds of times larger than the object, so this snapshot mimics what one might expect for plane-wave illumination and steady-state conditions. [(a) and (c)] TM-polarized SH field patterns for free charges only (a) and free and bound charges simultaneously (c). [(b) and (d)] Same as (a) and (c) for the TH fields. Qualitative differences are especially evident for the TH field, where bound charges tend to reinforce the main lobe that points directly backward. In all cases quantitative differences can be read on the side scales. Bound charges enhance SH conversion by a factor of ~2. TH generation is reduced by approximately the same factor. We have assumed $m^* = m_e$ and $n_0 = 5.8 \times 10^{22}/\text{cm}^3$. (Left) Depiction of the setup and pump magnetic field intensity profile recorded at the same time as the harmonics.

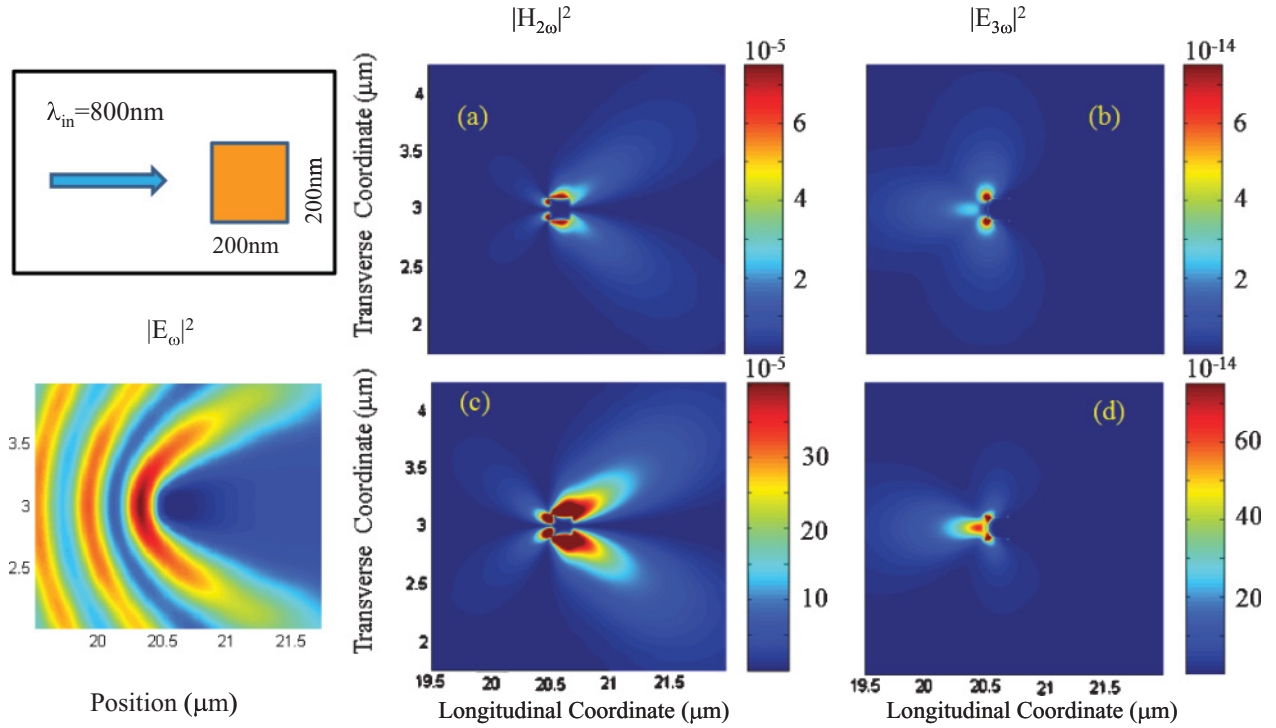

FIG. 8. (Color online) (Right) Same as Fig. 7, except for TE-polarized incident pump pulse. The snapshots are recorded when the peak of the pump pulse reaches the metal object. Unlike in Fig. 7, the generated SH signals are TM polarized (the H field is shown for convenience), while the TH fields are TE polarized (the E field is shown for convenience). The TM-polarized SH field patterns from free only (a) and from free and bound (c) charges. TE-polarized TH field patterns obtained from free only (b) and free and bound (e) charges. Qualitative differences are visible in both SH and TH patterns. The polarization of a resulting harmonic field under specific pumping conditions may be assessed by a full decomposition of the nonlinear sources in Eqs. (12) and (29). (Left) Depiction of the set up and electric field intensity profile recorded at the same time as the harmonics.

function and nonlinear contributions to harmonic generation may be significant if the fields become strongly confined to produce either (i) large evanescent wave vectors via strong field localization or (ii) large spatial derivatives near or just inside the metal surface. In the visible and IR ranges one can always count on the field penetrating inside the metal with consequent rapid spatial modulation. In what follows we concentrate on estimating perturbations to the linear dielectric constant in a grating of GaAs-filled, 60-nm-wide nanocavities carved on a silver substrate 100 nm thick. The grating is designed to be resonant at $\lambda = 1064$ nm, and displays enhanced transmission through a broad Fabry-Perot resonance. The structure is depicted in Fig. 9 along with the transmittance for TM-polarized light as a function of wavelength for a center-to-center distance (pitch) of ~570 nm between the nanocavities. The structure displays a maximum transmittance of ~280%, with field amplification values that range from 100 times for the transverse electric field intensity at the corners of the cavity to approximately 400 times for the magnetic field intensity. These relatively modest amplifications correspond to an effective cavity Q that hovers in the thousands and persists even for slightly rounded cavity corners.

The question relative to Eq. (13) and electron gas pressure may be formulated as follows: How large do pressure terms become relative to the linear driving field terms near resonance? We provide a partial answer in Fig. 10, where we compare the transverse and longitudinal pressure terms relative to their respective field components, when the peak of an incident 100-fs pulse reaches the cavity. The figure shows that the transverse pressure can alter the local, instantaneous transverse forcing term by approximately 1%–2%. Longitudinal pressure changes are strongest at the four corners of the cavity but amount to a more modest ~0.1%. Of course, these rates could increase or decrease depending on the exact magnitude of the Fermi velocity, tuning, and other geometrical factors. Nevertheless, these results show that electron gas pressure may act to either "soften" or "harden" the metal boundaries and to shift resonance conditions. Perhaps more significantly, while it is certainly important to establish the low-intensity bulk metal dielectric constant [68], these results point to the fact that in metallic, high-finesse, high-field cavities [55,67] it may be more important to determine and include terms not needed in more ordinary situations. Our estimates also indicate that in the case of metal layers or metal-dielectric multilayer stacks pressure contributions to the linear dielectric constant amount to at most one part in 10^5. We will report elsewhere more details on the dynamical effects of electron gas pressure, and the linear and nonlinear optical properties of metallic nanocavities in the enhanced transmission regime. Suffice it to say here that under extreme conditions effects due to electron gas pressure should at a minimum be assessed when strong field confinement occurs at the subwavelength scale, as it may account for resonance shifts on the order of tens of nanometers.

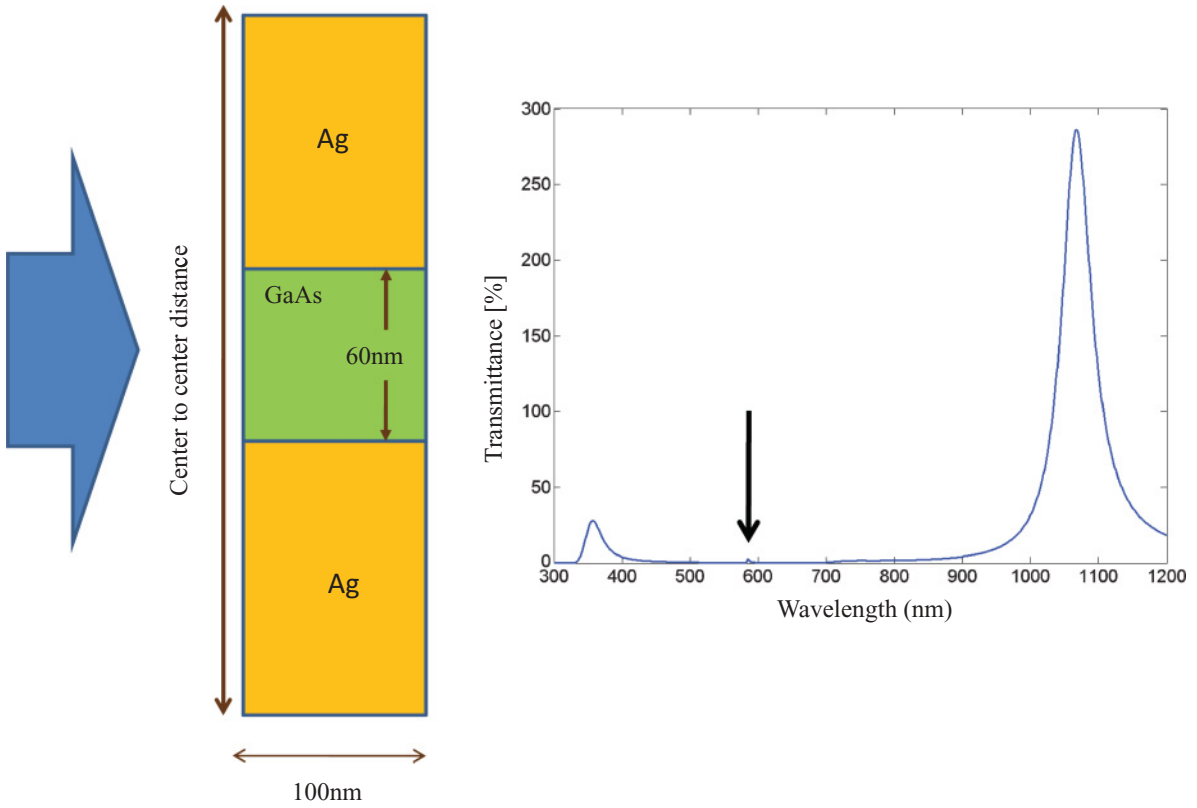

FIG. 9. (Color online) (Left) Wave incident on a GaAs-filled nanocavity 100 nm deep and 60 nm wide on an infinite grating. Departure from these parameters reduces transmittance and field amplification inside the cavity. (Right) Transmittance vs. wavelength, for a pitch of 570 nm, normalized to the energy that falls directly on the cavity area. Actual maximum total transmittance at the 1064-nm Fabry-Perot resonance is ~30%. A barely visible plasmonic transmission peak appears near 570 nm (indicated by the arrow).

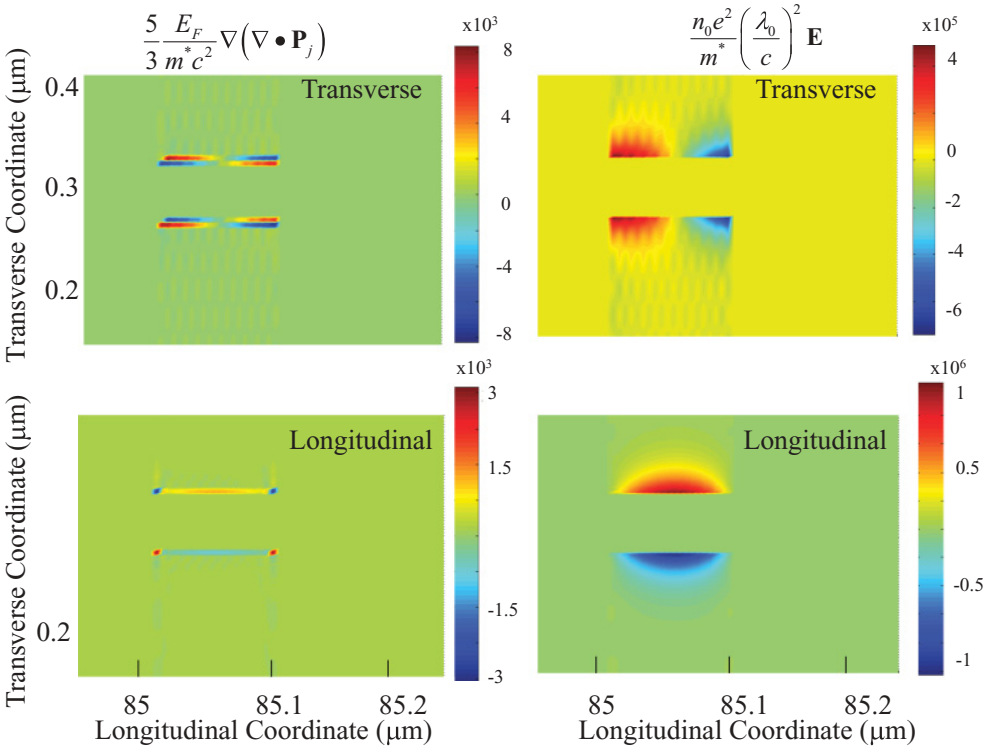

FIG. 10. (Color online) Assessing the effects of electron gas pressure. Top (bottom) Snapshot of transverse (longitudinal) pressure compared to its driving field term. The scales show that the transverse pressure is spatially modulated with amplitudes of order 1%–2% compared to the driving field term. The longitudinal pressure is modulated along its entire length but is strongest at the four corners of the cavity. Note: negative pressure simply indicates regions of pressure lower compared to contiguous regions.

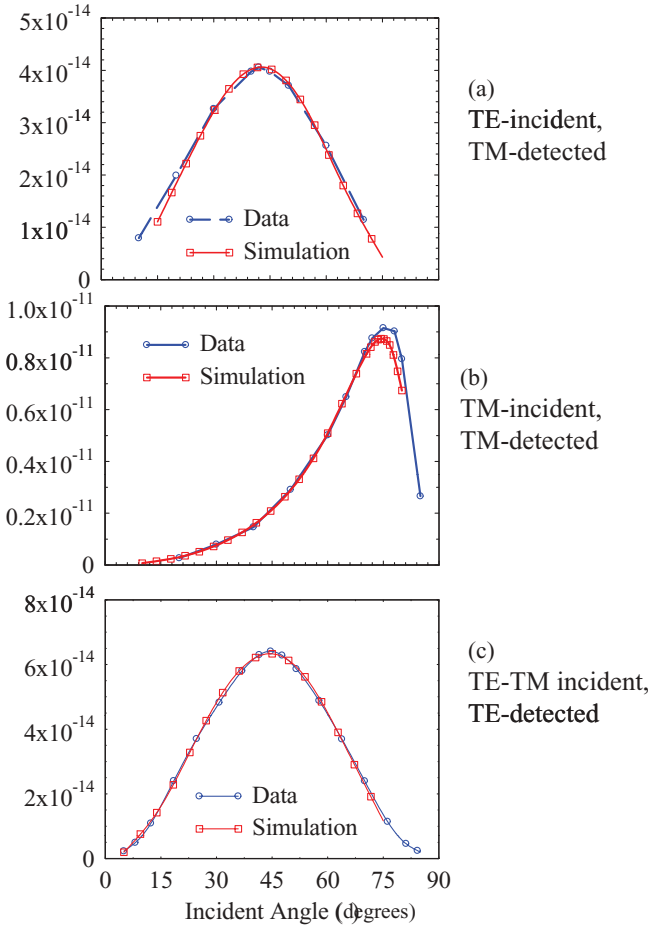

FIG. 11. (Color online) SH conversion efficiency under various pumping conditions. The data are taken from Ref. [71], where the average peak intensity is estimated at ~80 MW/cm^2. The shapes of theoretical curves agree well with the experimental data in all cases. If we choose $m^* = 0.5\,m_e$, and $n_0 = 3 \times 10^{22}/\text{cm}^3$, for both free and bound charges, excellent quantitative agreement emerges with the following choice of incident peak intensities: (a) ~60 MW/cm^2; (b) ~100 MW/cm^2; (c) ~300 MW/cm^2.

VIII. COMPARISON WITH EXPERIMENTAL RESULTS

We now attempt to make a comparison with experimental results [69–74] and choose Ref. [71] because it contains data for a pump field tuned at 1064 nm relative to silver for (i) TE-incident–TM-detected; (ii) TM-incident–TM-detected; and (iii) TE- and TM–incident–TE-detected polarization states. In Fig. 11 we show a comparison between our calculations and the data reported in Ref. [71]. The pulses used in Ref. [71] were 50 ps in duration, and the silver layer was 400 nm thick. Our simulations were carried out using 80-fs pulses and a silver layer 150 nm thick. As we mentioned earlier, dispersion plays a role only for pulses that are a few optical cycles in duration since the single metal layer has neither transmission nor reflection features. Substrate thickness becomes unimportant if the layer is substantially opaque. For simplicity, effective mass and density were chosen in order to maintain the ratio (i.e., plasma frequency) constant relative to the actual data. The input beam used in the experiment was slightly converging (ours was not), with a ~2-mm radius at the sample and a peak intensity of ~80 MW/cm^2. Once these parameters were fixed, Fig. 11(a) was fitted with an incident peak intensity of 60 MW/cm^2; in Fig. 11(b) we used ~100 MW/cm^2 and in Fig. 11(c) ~300 MW/cm^2. The comparison in Figs. 11(a) and 11(b) shows excellent quantitative and qualitative agreement between theory and experiment. Even though Fig. 11(c) also shows excellent qualitative agreement, our input intensity was approximately four times larger than that reported in Ref. [71]. Although these discrepancies are relatively small and could be easily accounted for by modest power and beam radius fluctuations and perhaps even the focusing geometry, more stringent experimental and theoretical tests are under way to further validate the model under a variety of conditions.

IX. CONCLUSIONS

We have presented a model to study propagation phenomena and harmonic generation in nanostructures in the ultrashort pulse regime, with the inclusion of linear and nonlinear effects due to free and bound charges simultaneously. Free electrons are modeled using the hydrodynamics model, comprising effects due to convection and electron gas pressure. Bound charges are modeled by Lorentz oscillators under the action of electric, magnetic, and nonlinear restoring forces. We have shown that the influence of bound charges and convection on harmonic generation can be qualitatively and quantitatively appreciable. In addition, we have briefly discussed dynamical changes that can occur in resonant metal gratings or nanocavities as a result of electron gas pressure. Although we have derived explicit nonlinear electron gas pressure terms, we have not discussed their influence on harmonic generation. We will do so in a different setting, in the more specific context of GaAs-filled resonant nanocavities. Finally, we found our calculations to be in good qualitative and quantitative agreement with the data of Ref. [71], where SHG from a silver film was reported under different incident pumping conditions. Although some differences remain and more tests cases are actively under consideration, our calculations and comparisons so far show that this approach may suffice to explain all relevant features of harmonic generation, provided the model includes all crucial dynamical aspects, namely convection, electron gas pressure, and contributions from core electrons. Previous experiments do show that it is possible for both the effective electron mass and density to experience relatively large fluctuations due to stresses in the fabrication system. The values that we used appear to yield reflection second-harmonic conversion efficiencies and curve shapes that are almost identical to experimental results, under three very different pumping conditions. These values could be refined by seeking SHG and THG data for transmitted fields, say, for thin metal layers or in the kind of transparent multilayer structure we have described. Calibration of transmission curves could be used to narrow down mass and density values further. However, one should not lose sight of the fact that whatever values one uses are bound to be system specific and may easily change for different experimental conditions and setups.

ACKNOWLEDGMENTS

We thank Joseph Haus and Henry Everitt for useful discussions and suggestions. V.R. and M.C. thank the Army Research Lab, London Office, for partial financial support. N.A. thanks the National Research Council for financial support.

[1] P. S. Pershan, Phys. Rev. **130**, 919 (1963).
[2] N. Bloembergen, Proc. IEEE **51**, 124 (1963).
[3] N. Bloembergen, R. K. Chang, S. S. Jha, and C. H. Lee, Phys. Rev. **174**(3), 813 (1968).
[4] E. Adler, Phys. Rev. **134**, A728 (1964).
[5] S. Jha, Phys. Rev. Lett. **15**, 412 (1965).
[6] S. Jha, Phys. Rev. **140**, A2020 (1965).
[7] S. S. Jha and C. S. Warke, Phys. Rev. **153**, 751 (1967).
[8] F. Brown, R. E. Parks, and A. M. Sleeper, Phys. Rev. Lett. **14**, 1029 (1965).
[9] F. Brown and R. E. Parks, Phys. Rev. Lett. **16**, 507 (1966).
[10] N. Bloembergen and Y. R. Shen, Phys. Rev. **141**, 298 (1966).
[11] H. Sonnenberg and H. Heffner, J. Opt. Soc. Am. **58**, 209 (1968).
[12] N. Bloembergen and C. H. Lee, Phys. Rev. Lett. **19**, 835 (1967).
[13] J. Rudnick and E. A. Stern, Phys. Rev. B **4**, 4274 (1971).
[14] C. C. Wang and A. N. Duminski, Phys. Rev. Lett. **20**, 668 (1968).
[15] J. C. Quail and H. J. Simon, Phys. Rev. B **31**, 4900 (1985).
[16] J. B. Pendry, Phys. Rev. Lett. **85**, 3966 (2000).
[17] T. W. Ebessen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, Nature (London) **391**, 667 (1998).
[18] M. W. Klein, M. Wegener, N. Feth, and S. Linden, Opt. Express **15**, 5238 (2007).
[19] Y. Zeng, W. Hoyer, J. Liu, S. W. Koch, and J. V. Moloney, Phys. Rev. B **79**, 235109 (2009).
[20] 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).
[21] 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).
[22] J. Butet, J. Duboisset, G. Bachelier, I. Russier-Antoine, E. Benichou, C. Jonin, and P. F. Brevet Nano Lett. **10**, 1717 (2010).
[23] V. K. Valev, A. V. Silhanek, N. Verellen, W. Gillijns, P. Van Dorpe, O. A. Aktsipetrov, and G. A. E. Vandenbosch, Phys. Rev. Lett. **104**, 127401 (2010).
[24] M. Airola, Y. Liu, and S. Blair, J. Opt. A: Pure Appl. Opt. **7**, S118 (2005).
[25] Tingjun Xu, Xiaojin Jiao, and Steve Blair, Opt. Express **17**, 23582 (2009).
[26] 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).
[27] A. Lesuffler, L. Kiran Swaroop Kumar, and R. Gordon, Appl. Phys. Lett. **88**, 261104 (2006).
[28] W. Fan, S. Zhang, K. J. Malloy, S. R. J. Brueck, N. C. Panoiu, and R. M. Osgood, Opt. Express **14**, 9570 (2006).
[29] M. A. Vincenti, M. De Sario, V. Petruzzelli, A. D'Orazio, F. Prudenzano, D. de Ceglia, N. Akozbek, M. J. Bloemer, P. Ashley, and M. Scalora, Proc. SPIE **6987**, 69870O (2008).
[30] 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).
[31] G. D'Aguanno, M. C. Larciprete, N. Mattiucci, A. Belardini, M. J. Bloemer, E. Fazio, O. Buganov, M. Centini, and C. Sibilia, Phys. Rev. A **81**, 013834 (2010).
[32] W. L. Schaich and B. S. Mendoza, Phys. Rev. B **45**, 14279 (1992).
[33] C. D. Hu, Phys. Rev. B **40**, 7520 (1989).
[34] A. Liebsch, Phys. Rev. Lett. **61**, 1233 (1988).
[35] D. Krause, C. W. Teplin, and C. T. Rogers, J. Appl. Phys. **96**, 3626 (2004).
[36] F. Xiang Wang, F. J. Rodríguez, W. M. Albers, R. Ahorinta, J. E. Sipe, and M. Kauranen, Phys. Rev. B **80**, 233402 (2009).
[37] Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 2002).
[38] D. Maystre, M. Neviere, and R. Reinish, Appl. Phys. A **39**, 115 (1986).
[39] J. L. Coutaz, D. Maystre, M. Neviere, and R. Reinisch, J. Appl. Phys. **62**, 1529 (1987).
[40] 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).
[41] M. Corvi and W. L. Schaich, Phys. Rev. B **33**, 3688 (1986).
[42] A. Eguiluz and J. J. Quinn, Phys. Rev. B **14**, 1347 (1976).
[43] A. Sekiyama *et al.* [http://arxiv.org/abs/1003.3294].
[44] G. Petrocelli, S. Martelucci, and R. Francini, Appl. Phys. A **56**, 263 (1993).
[45] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, London/New York, 1985), p. 350.
[46] H. Ehrenreich and H. R. Philipp, Phys. Rev. **128**, 1622 (1962).
[47] H. R. Philipp and H. Ehrenreich, Phys. Rev. **129**, 1550 (1963).
[48] B. R. Cooper, H. Ehrenreich, and H. R. Philipp, Phys. Rev. **138**, A494 (1965).
[49] P. Guyot-Sionnest and Y. R. Shen, Phys. Rev. B **38**, 7985 (1988).
[50] A. Benedetti, M. Centini, C. Sibilia, and M. Bertolotti, J. Opt. Soc. Am. B **27**, 408 (2010).
[51] M. del Castillo-Mussot, G. J. Vazquez Fonzeca, and G. H. Cocoletzi, Rev. Mex. Fis. **49**, 231 (2003).
[52] E. L. Linder, Phys. Rev. **49**, 753 (1936).
[53] G. Manfredi and F. Haas, Phys. Rev. B **64**, 075316 (2001).
[54] N. Crouseilles, P. A. Hervieux, and G. Manfredi, Phys. Rev. B **78**, 155412 (2008).
[55] M. Neviere and R. Reinisch, Phys. Rev. B **26**, 5403 (1982).
[56] M. A. Vincenti, D. de Ceglia, M. Buncick, N. Akozbek, M. J. Bloemer, and M. Scalora, J. Appl. Phys. **107**, 053101 (2010).
[57] A. T. Georges and N. E. Karatzas, Appl. Phys. B: Lasers and Optics **81**, 479 (2005).
[58] N. A. Papadogiannis, P. A. Loukakos, and S. D. Moustaizis, Opt. Commun. **166**, 133 (1999).
[59] H. Teng and Chunlei Guo, Appl. Phys. Lett. **85**, 1110 (2004).
[60] M. Scalora and M. E. Crenshaw, Opt. Commun. **108**, 191 (1994).
[61] A. G. Mathewson, H. Aronsson, and L. G. Bernland, J. Phys. F: Metal Phys. **2**, L39 (1972).

[62] J. Szczyrbowski, J. Phys. D: Appl. Phys. **19**, 1257 (1986).
[63] G. Grimvall, Phys. Kondens. Mater. **11**, 279 (1970).
[64] J. Dryzek and A. Czapla, J. Mater. Sci. Lett. **4**, 154 (1985).
[65] A. I. Maaroof, A. Gentle, G. B. Smith, and M. B. Cortie, J. Phys. D: Appl. Phys. **40**, 5675 (2009).
[66] D. J. Nash and J. R. Sambles, J. Mod. Opt. **46**, 1793 (1999).
[67] M. I. Stockman, D. J. Bergman, C. Anceau, S. Brasselet, and J. Zyss, Phys. Rev. Lett. **92**, 057402 (2005).
[68] V. P. Drachev, U. K. Chettiar, A. V. Kildishev, H.-K. Yuan, W. Cai, and Vladimir M. Shalaev, Opt. Express **16**, 1186 (2006).
[69] C. Hubert, L. Billot, P.-M. Adam, R. Bachelot, P. Royer, J. Grand, D. Gindre, K. D. Dorkenoo, and A. Fort, Appl. Phys. Lett. **90**, 181105 (2007).
[70] N. N. Akhmediev, I. V. Mel'nikov, and L. J. Robur, Laser Phys. **4**, 1194 (1994).
[71] K. A. O'Donnell and R. Torre, New J. Phys. **7**, 154 (2005).
[72] C. S. Chang and J. T. Lue, Surf. Sci. **393**, 231 (1997).
[73] R. Vollmer, M. Straub, and J. Kirshner, Surf. Sci. **352-354**, 684 (1996).
[74] C. C. Tzeng and J. T. Lue, Phys. Rev. A **39**, 191 (1989).