Politecnico di Bari

Repository Istituzionale dei Prodotti della Ricerca del Politecnico di Bari

Nonlinear Duffing oscillator model for third harmonic generation

(Article begins on next page)

13 May 2024

Nonlinear Duffing Oscillator Model for Third Harmonic Generation

M. Scalora[1], M. A. Vincenti[2], D. de Ceglia[2], C. M. Cojocaru[3], M. Grande[4], J. W. Haus[5]

[1] *Charles M. Bowden Research Center, AMRDEC, RDECOM, Redstone Arsenal, AL 35898-5000*

[2] *National Research Council - AMRDEC, Charles M. Bowden Research Center, Redstone Arsenal, AL 35898*

[3] *Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, 08222 Terrassa, Spain*

[4] *Dipartimento di Ingegneria Elettrica e dell'Informazione, Politecnico di Bari, Via Re David 200, 70125 Bari, Italy*

[5] *Electro-Optics Program, University of Dayton, Dayton, OH 45469-2951*

ABSTRACT

We employ a classical, nonlinear Lorentz-Duffing oscillator model to predict third harmonic conversion efficiencies in the ultrafast regime, from a variety of metal nanostructures, including smooth, isolated metal layers, a metal-dielectric photonic band gap structure, and a metal grating. As expected, the plasmonic grating yields the largest narrow-band conversion efficiencies. However, interference phenomena at play within the multilayer stack yield comparable, broad-band conversion. The method includes both linear and nonlinear material dispersions that in turn sensitively depend on linear oscillator parameters. Concurrently, and unlike other techniques, the integration scheme is numerically stable. By design, one thus avoids the introduction of explicit, third-order nonlinear coefficients and also takes into account linear and nonlinear material dispersions simultaneously, elements that are often necessary to fully understand many of the subtleties of the interaction of light with matter.

1. Introduction

The classical, Lorentz oscillator model of atomic media is an immensely pedagogical tool that serves to represent most linear and nonlinear optical phenomena [1]. As such, it is also a natural springboard for the description and understanding of many quantum phenomena. In this paper we use the nonlinear Lorentz-Duffing model to examine primarily third harmonic generation (THG) from several geometrical arrangements that include relatively thin, isolated layers of metal, a metal-dielectric photonic band gap structure, and a metal grating, however, without neglecting second harmonic generation (SHG). In centrosymmetric materials like metals, SHG is most often examined as the direct product of the interaction of an incident field with free electrons only, with secondary contributions of bound electrons applied only to the linear dielectric constant. In reality, just as the linear dielectric constant is affected by both free electrons and interband transitions from electrons in the valence band, SHG can also arise from both conduction and inner-core electrons [2], due to a combination of spatial symmetry breaking (the mere presence of interfaces), the magnetic portion of the Lorentz force, to a lesser extent the interaction of third harmonic and pump photons (down-conversion), and other effective nonlinearities induced by quantum tunneling mechanisms [3-5] if metal components are in close proximity. By the same token, the third-order nonlinearity, $\chi^{(3)}$, arising from bound charges generates most of the third harmonic signal, subject to screening due to a free-electron spill-out effect and other geometrical considerations [5, 6]. To a much smaller degree the interaction of pump and second harmonic photons also produces cascaded THG. Therefore, in our approach the classical, nonlinear oscillator model will be pivotal in these systems, where a combination of free (Drude) and bound (Lorentz) electrons generally suffices to describe second and third order processes.

In general, the nonlinear polarization of a medium may be written according to a well-established, ascending hierarchy of electric and magnetic dipoles and multipoles [7, 8]. In practical terms this means explicit introduction of a nonlinear polarization that requires *a priori* knowledge of second, third and higher order nonlinear coefficients, in either scalar or tensorial form. On the other hand, the nonlinear Lorentz-Duffing oscillator model is a more natural, possibly more instructive starting point that implicitly accounts for linear and nonlinear material dispersion [1], that turns the polarization into an independent parameter and is able to describe an entire class of nonlinear optical phenomena, including, perhaps most notably, intrinsic (or mirror-less) optical bistability [9]. A charge e is under the action of a number of linear and nonlinear forces that tend to shift its position as it orbits around the nucleus. Accordingly, an imbalance between internal and external forces causes the electron to depart from its equilibrium position. As an example, for bound electrons that occupy a given orbital in a centrosymmetric material like silver, Newton's second law may be written as:

$$m^*\ddot{\mathbf{r}}(t) + \gamma\, m^*\dot{\mathbf{r}}(t) + k\mathbf{r}(t) - m^*b\big(\mathbf{r}(t)\bullet\mathbf{r}(t)\big)\mathbf{r}(t) = e\mathbf{E}(\mathbf{r},t) + \frac{e}{c}\dot{\mathbf{r}}\times\mathbf{H}(\mathbf{r},t)\,,\tag{1}$$

where $\mathbf{r}(t)$ is the displacement from orbital equilibrium; $\mathbf{E}, \mathbf{H}$ are the applied electric and magnetic fields, respectively; m^* is the electron's effective mass; k is the spring constant associated with a linear restoring force; b is a coefficient associated with a nonlinear, third-order restoring force; c is the speed of light in vacuum; γ is the damping coefficient. As we will see later, the motion of nearly-free, conduction electrons is accounted for by supplementing Eq.(1) with an equation similar to Eq.(1), but without linear and nonlinear spring restoring forces.

Before delving into the complex aspects of calculating conversion efficiencies for specific, yet emblematic geometries, it is worthwhile recalling some of the most salient points

3

that relate to the form of Eq.(1), specifically regarding nonlinear material dispersion. An alternative discussion of the derivation of nonlinear material dispersion based on the nonlinear oscillator model may be found in references [1, 10, 11]. For clarity, we temporarily neglect SHG but will reintroduce it later in the general treatment of second and third order processes in centrosymmetric materials. For THG, incident pump photons of frequency ω generate photons having three times the pump frequency. Therefore, it is reasonable to expect that both pump and third harmonic photons are present simultaneously. As a result, at the electron position the total field may be thought of as a superposition of fundamental frequency (FF) and third harmonic (TH) fields as follows:

$$
\begin{aligned}
\mathbf{E} &= \mathbf{E}_{\omega} e^{-i\omega t} + \mathbf{E}_{3\omega} e^{-3i\omega t} + c.c. \\
\mathbf{H} &= \mathbf{H}_{\omega} e^{-i\omega t} + \mathbf{H}_{3\omega} e^{-3i\omega t} + c.c.
\end{aligned}
\tag{2}
$$

Here $\mathbf{H}_{\omega}, \mathbf{E}_{\omega}, \mathbf{H}_{3\omega}, \mathbf{E}_{3\omega}$ are complex field envelopes that for the moment are assumed to be constant in time, an assumption that we will later discard. By the same token, a simplified solution for the electron displacement from its equilibrium position may also be written as a superposition of FF and TH amplitudes, so that:

$$
\mathbf{r} = \mathbf{r}_{\omega} e^{-i\omega t} + \mathbf{r}_{3\omega} e^{-3i\omega t} + c.c..
\tag{3}
$$

Substituting Eqs.(2-3) into Eq.(1), retaining lowest order terms, neglecting the nonlinear, magnetic term, and equating terms that oscillate at the same frequency, one obtains:

$$
\begin{aligned}
\mathbf{r}_{\omega} &\approx \frac{e}{m^{*}(\omega_0^2 - \omega^2 - i\gamma\,\omega)}\mathbf{E}_{\omega} + \frac{b}{(\omega_0^2 - \omega^2 - i\gamma\omega)}\,|\,\mathbf{r}_{\omega}\,|^2\,\mathbf{r}_{\omega} \\
\mathbf{r}_{3\omega} &\approx \frac{e}{m^{*}\left(\omega_0^2 - 9\omega^2 - 3i\gamma\,\omega\right)}\mathbf{E}_{3\omega} + \frac{b}{(\omega_0^2 - 9\omega^2 - 3i\gamma\,\omega)}\left(\mathbf{r}_{\omega} \bullet \mathbf{r}_{\omega}\right)\mathbf{r}_{\omega}
\end{aligned}
\tag{4}
$$

4

where $\omega_0^2 = \dfrac{k}{m^*}$ is the oscillator's resonance frequency. In turn, approximate solutions may now be pursued for the amplitudes $\mathbf{r}_\omega$ and $\mathbf{r}_{3\omega}$. For instance, the first of Eqs.(4) becomes:

$$\mathbf{r}_\omega = \frac{e}{m^*(\omega_0^2 - \omega^2 - i\gamma\omega)\left(1 - \dfrac{b}{(\omega_0^2 - \omega^2 - i\gamma\omega)}|\mathbf{r}_\omega|^2\right)}\mathbf{E}_\omega \qquad . \tag{5}$$

If the field is not too intense, and/or if tuning is not too close to resonance such that the condition $\left|b|\mathbf{r}_\omega|^2/(\omega_0^2 - \omega^2 - i\gamma\omega)\right| << 1$ is satisfied at all times, a stipulation not required in an approach that includes the integration of Eq.(1), the denominator of Eq.(5) may be expanded so that:

$$\mathbf{r}_\omega = \frac{e}{m^*(\omega_0^2 - \omega^2 - i\gamma\omega)}\mathbf{E}_\omega\left(1 + \frac{b}{(\omega_0^2 - \omega^2 - i\gamma\omega)}|\mathbf{r}_\omega|^2 + ...\right) \qquad . \tag{6}$$

To first order in the field amplitude we have:

$$\mathbf{r}_\omega = \frac{e}{m^*(\omega_0^2 - \omega^2 - i\gamma\omega)}\mathbf{E}_\omega + \frac{be^3}{m^{*3}(\omega_0^2 - \omega^2 - i\gamma\omega)^3(\omega_0^2 - \omega^2 + i\gamma\omega)}|\mathbf{E}_\omega|^2\,\mathbf{E}_\omega. \tag{7}$$

Eq.(7) is then used to solve for the second of Eq.(4). Upon substitution we find:

$$\mathbf{r}_{3\omega} = \frac{e}{m^*\left(\omega_0^2 - 9\omega^2 - 3i\gamma\omega\right)}\mathbf{E}_{3\omega} + \frac{be^3}{m^{*3}(\omega_0^2 - 9\omega^2 - 3i\gamma\omega)(\omega_0^2 - \omega^2 - i\gamma\omega)^3}\left(\mathbf{E}_\omega \bullet \mathbf{E}_\omega\right)\mathbf{E}_\omega \tag{8}$$

Recognizing that $\mathbf{p}_{\omega,3\omega} = e\mathbf{r}_{\omega,3\omega}$ is the dipole moment, and that $\mathbf{P}_{\omega,3\omega} = Ne\mathbf{r}_{\omega,3\omega}$ is the total polarization density per unit volume (N is the oscillator density), we may write:

$$\mathbf{P}_\omega = \chi_\omega^{(1)}\mathbf{E}_\omega + \chi_\omega^{(3)}|\mathbf{E}_\omega|^2\,\mathbf{E}_\omega \qquad , \tag{9a}$$

and

$$\mathbf{P}_{3\omega} = \chi_{3\omega}^{(1)}\mathbf{E}_\omega + \chi_{3\omega}^{(3)}\left(\mathbf{E}_\omega \bullet \mathbf{E}_\omega\right)\mathbf{E}_\omega \tag{9b}$$

where

$$\chi_\omega^{(1)} = \frac{Ne^2}{m^*(\omega_0^2 - \omega^2 - i\gamma\omega)}; \qquad \chi_\omega^{(3)} = \frac{Nbe^4}{m^{*3}(\omega_0^2 - \omega^2 - i\gamma\omega)^3(\omega_0^2 - \omega^2 + i\gamma\omega)} \qquad ; \qquad (10)$$

and

$$\chi_{3\omega}^{(3)} = \frac{Nbe^4}{m^{*3}(\omega_0^2 - \omega^2 - i\gamma\omega)^3(\omega_0^2 - 9\omega^2 - 3i\gamma\omega)} \quad . \qquad (11)$$

Put another way, the solution of Eq.(1) implicitly contains linear and nonlinear material dispersions [1] that are predetermined by the choice of linear material parameters like density, plasma frequency, damping rates, effective mass, and resonance frequency, at least in a context where the approximations used to arrive at Eqs.(6-8) remain valid.

Although an approach using Eq.(1) is often mentioned as desirable due to the apparently general, comprehensive nature of the solution, a technique to solve Eq.(1) is practically never employed because typical integration schemes of Eq.(1) based on a straightforward finite difference approach are deemed unstable [12, 13], unless special precautions are taken. For example, the authors of Refs.[12, 13] suggest that the cubic nonlinearity should be "balanced" by the introduction of an artificial, nonlinear quintic term characterized by a coefficient having sign opposite to that of the cubic term, in order to provide a counter-weight that removes runaway solutions. Therefore, the usual tactic is to avoid direct integration of Eq.(1), and opt instead for the use of a system of steady-state Bloch equations that describe either a two- or a multi-level atom tuned far from resonance in order to emulate a classical oscillator [14]. This approach allows the derivation of a standard, nonlinear saturable absorption term that is expanded perturbatively in ascending powers of the electric field intensity [12, 13], similar to what one obtains by expanding Eq.(5), thus justifying the use of saturable absorption in the nonlinear Duffing oscillator model given by Eq.(1) [15, 16]. Below we will present an alternative,

numerically stable approach that combines a spectral, fast Fourier transform (FFT), time-domain beam propagation method to solve Maxwell's equation and a modified predictor-corrector technique to solve Eq.(1).

The nature and the size of the coefficient b in Eqs. (10-11) remains to be ascertained. One usually argues that the nonlinear restoring force becomes important when it is of the same order of magnitude as the linear restoring force, i.e. $\omega_0^2 |\mathbf{r}_0| \approx b |\mathbf{r}_0^3|$, which yields $b \approx \dfrac{\omega_0^2}{|\mathbf{r}_0^2|}$ [1]. Ordinarily, for dielectric materials the value of $|\mathbf{r}_0|$ is chosen to coincide with the atomic diameter, or approximately 3Å, which simultaneously represents the lattice constant and the maximum possible excursion of a valence electron orbiting the atom. For noble metals, $|\mathbf{r}_0|$ should probably be chosen to correspond to the size of the given orbital of provenance, (for example, $|\mathbf{r}_0| \sim 0.5$Å for 4-d shell electrons in silver, and 5d-shell in gold) which may be ascertained by calculating the electronic wave functions [17, 18]. One possibility may be to obtain an optimized value of b, given a known estimate of $\chi^{(3)}(\omega = 0)$ [19]. Alternatively, one may adopt a nonlinear polarization model of the usual type, $\mathbf{P}^{NL} = \chi^{(3)} \mathbf{E}^3$, as we will do later, and retrieve a generic $\chi^{(3)}$ value that yields the same conversion efficiency as does the choice $b \approx \dfrac{\omega_0^2}{|\mathbf{r}_0^2|}$. Nevertheless, the combination of: (i) imprecise knowledge of the electron's effective mass in the given orbital, which presumably has more inertia compared to nearly-free conduction electrons; (ii) the fact that at high frequencies more than one bound electron per atom may be available [20]; and (iii) as yet unaccounted-for oscillator strengths and structure-related geometrical factors make the recovery of accurate $\chi^{(3)}$ values somewhat arduous.

7

We now consider an example. In Fig.1 (a) we plot the dielectric data of silver, as reported in Palik's handbook [21], along with a possible fit using one Drude and *two* Lorentz oscillators, as follows:

$$\varepsilon(\tilde{\omega}) = 1 - \frac{\tilde{\omega}_{pf}^2}{\tilde{\omega}^2 + i\tilde{\gamma}_f\tilde{\omega}} - \frac{\tilde{\omega}_{p1}^2}{\tilde{\omega}^2 - \tilde{\omega}_{01}^2 + i\tilde{\gamma}_{01}\tilde{\omega}} - \frac{\tilde{\omega}_{p2}^2}{\tilde{\omega}^2 - \tilde{\omega}_{02}^2 + i\tilde{\gamma}_{02}\tilde{\omega}}. \tag{12}$$

Here $\tilde{\omega} = 1/\lambda$, where λ is in microns, $\tilde{\gamma}_f = 0.0573, \tilde{\gamma}_{01} = \tilde{\gamma}_{02} = 1.42, \tilde{\omega}_{pf} = 6.965$, $\tilde{\omega}_{p1} = 4.25$, $\tilde{\omega}_{p2} = 5.5$, $\tilde{\omega}_{01} = 3.75$ $\tilde{\omega}_{02} = 5$. A comparison of the curves reveals reasonably good agreement
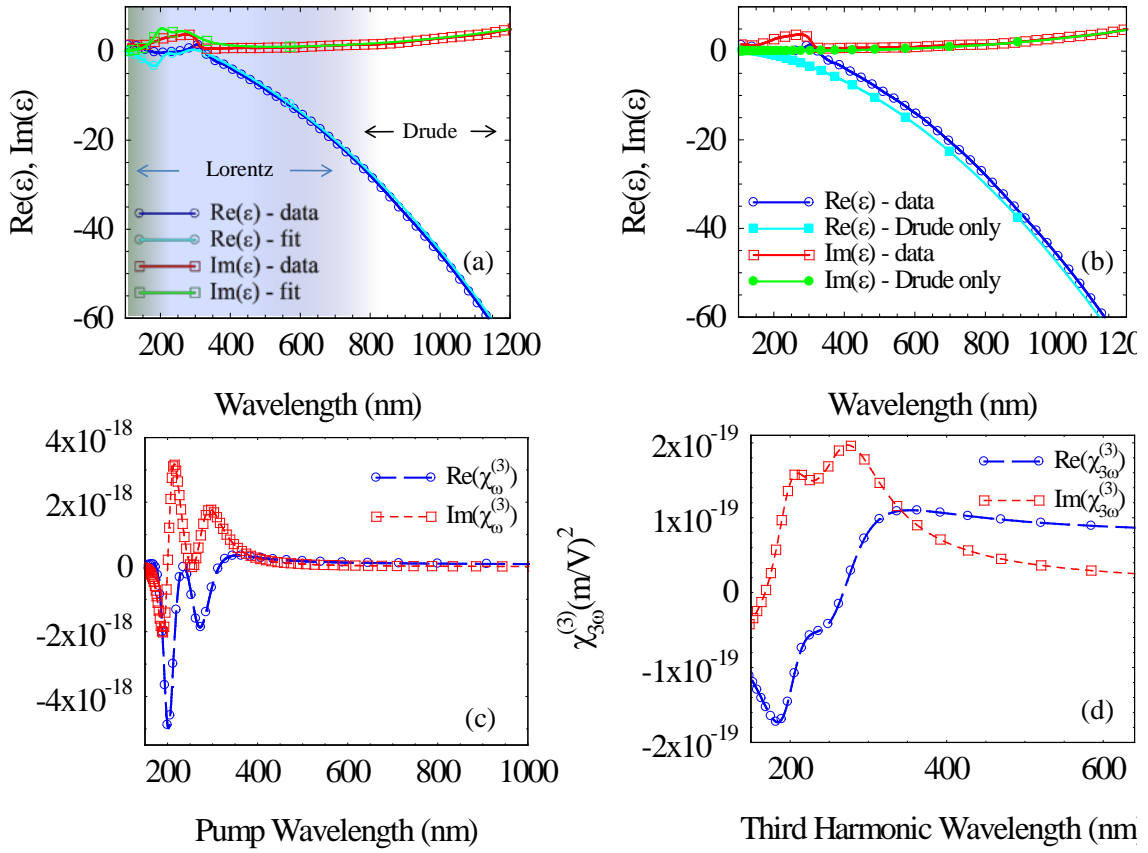

Fig.1: (a) Bulk silver data and the function in Eq.(12). (b) Palik's data compared to a Drude-only contribution. Both figures (a) and (b) suggest the importance of the Lorentz contributions at short wavelengths. Nonlinear susceptibility for self-phase modulation and nonlinear absorption (c), and third harmonic generation (d), obtained from Eqs.(10) and (11), respectively. The nonlinear dispersion curves are very sensitive to the linear fit and the effective mass m^*. Using Eq.(12) we estimate $m_{b1}^* \approx 98m_e$ and $m_{b2}^* \approx 58m_e$. The label "Lorentz" in (a) is to be understood as connoting a smooth transition region where the bound oscillators are increasingly important.

over most of the range illustrated (from ~150nm to ~1200nm). The increasingly darker shading in Fig. 1(a) connotes progressively greater bound electron contributions. In Fig. 1(b) we plot the dielectric constant in Eq.(12) *without* the Lorentz oscillator portions. The plot clearly shows that bound electrons contribute significantly and become increasingly important to the dielectric constant beginning in the near-IR, as the shading in Fig. 1(a) is meant to illustrate. In Figs.1 (c) and (d) we show the nonlinear dispersion curves obtained from the dielectric function in Eq.(12) using Eqs.(10-11). Looking ahead at the scaled equations of motion Eqs.(25-26), as an illustration of the approach, we may use Eq.(12) to estimate the effective masses for each of the two Lorentz oscillators using the definitions of the scaled plasma frequencies, $\tilde{\omega}_{pj}^2 = 4\pi n_{0j} e^2 \lambda_r^2 / (m_{bj}^* c^2)$. For example, given that $\tilde{\omega}_{p1}^2 = 4\pi n_{01} e^2 \lambda_r^2 / (m_{b1}^* c^2) = 4.25^2$ and $\tilde{\omega}_{p2}^2 = 4\pi n_{02} e^2 \lambda_r^2 / (m_{b2}^* c^2) = 5.5^2$, with $\lambda_r = 1\mu m$ and $n_{01} = n_{02} = 5 \times 10^{22} / cm^3$, we have $m_{b1}^* \approx 98 m_e$ and $m_{b2}^* \approx 58 m_e$. Then, assuming $\omega_0 \approx 8 \times 10^{15} \sec^{-1}$ and $r_0 = 0.5\text{Å}$, both $\chi_\omega^{(3)}$ and $\chi_{3\omega}^{(3)}$ are of order $10^{-19} (m/V)^2$ in the relatively flat spectral regions $\lambda > 500nm$. Near resonance, $\chi_\omega^{(3)}$ increases by two orders of magnitude.

The sensitivity of the nonlinear dispersion curves in Fig.1 may be ascertained by limiting the wavelength range of interest, for example, down to 500nm, and by dropping the second Lorentz oscillator in Eq.(12). In that case, one would find that despite continued agreement between the linear data and Eq.(12) *minus the second oscillator*, the amplitudes of the *nonlinear* dispersion curves can change by as much as one order of magnitude, significantly impacting the predicted conversion efficiencies. On the other hand, adding one more Lorentz oscillator to the two already in place in Eq.(12), to attempt to fit the data beyond 150nm, causes minimal changes to the nonlinear dispersion curves of Fig.1. Therefore, how one approaches the problem is of particular

9

importance in view of the fact that the reported values of $\chi^{(3)}$ may differ by several orders of magnitude, depending on pulse duration and the structure investigated [22]. We note that our approach does not include a description of hot electrons temporarily elevated from the valence to the conduction band, a process usually handled by introducing a two-temperature model [22, and references therein]. Here we assume these effects remain latent through the use of low-repetition rate, femtosecond pulses. The extension of the model to include thermal effects is easily accomplished, will not be treated here, but will be the subject of later refinements of the model.

2. Method of Solution

Before illustrating the full nonlinear problem we outline the method of solution by examining the propagation of a field coupled to a single Lorentz oscillator that contains a generic nonlinear term $\mathbf{P}^{NL}$. Neglecting effects due to the magnetic portion of the Lorentz force, we have:

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}$$
$$\nabla \times \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\frac{\partial \mathbf{P}}{\partial t} \qquad . \tag{13}$$
$$\ddot{\mathbf{P}} + \gamma\dot{\mathbf{P}} + \omega_0^2\mathbf{P} + \mathbf{P}^{NL} = \frac{Ne^2}{m}\mathbf{E}$$

As noted previously, the stability of the integration scheme of Eqs.(13) is an issue if $\mathbf{P}^{NL} \sim (\mathbf{P}\bullet\mathbf{P})\mathbf{P}$, unless either saturable absorption or artificial, higher-order contributions are introduced. Our approach is based on calculating the fields' spatial derivatives with high accuracy using FFTs, while material equations are integrated using a modified, second-order-accurate, predictor-corrector method. Without loss of generality, we assume the incident field is TM-polarized (Fig. 2), and decompose the **E** and **H** fields as products of generic, complex envelope functions and terms that contain free-space carrier wave-vector and angular frequency

10

that correspond to initial conditions for a pulse located in free space approaching the structure. The preservation of all spatial and temporal derivatives accounts for dynamical changes to the instantaneous phases and amplitudes of the fields. The **E** field is polarized on the y-z plane, the **H** field is polarized along the x-direction, as in Fig.2, and may be written as follows:

$$\mathbf{E} = \begin{pmatrix} E_{\hat{y}} \\ E_{\hat{z}} \end{pmatrix} = \begin{pmatrix} \mathbf{j}\left(E_{TM\hat{y}}^{\omega}(\mathbf{r},t)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + E_{TM\hat{y}}^{2\omega}(\mathbf{r},t)e^{2i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + E_{TM\hat{y}}^{3\omega}(\mathbf{r},t)e^{3i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + c.c. \right) \\ \mathbf{k}\left(E_{TM\hat{z}}^{\omega}(\mathbf{r},t)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + E_{TM\hat{z}}^{2\omega}(\mathbf{r},t)e^{2i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + E_{TM\hat{z}}^{3\omega}(\mathbf{r},t)e^{3i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + c.c. \right) \end{pmatrix}, \quad (14)$$

$$\mathbf{H} = H_{\hat{x}} = \mathbf{i}\left(H_{TM\hat{x}}^{\omega}(\mathbf{r},t)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + H_{TM\hat{x}}^{2\omega}(\mathbf{r},t)e^{2i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + H_{TM\hat{x}}^{3\omega}(\mathbf{r},t)e^{3i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + c.c. \right). \quad (15)$$

Then, for each harmonic component, the field, current, and polarization envelopes obey coupled equations of motion of the type:

$$\frac{\partial H_{\hat{x}}}{\partial \tau} = i\beta\left(H_{\hat{x}} + E_{\hat{z}}\sin\theta_i + E_{\hat{y}}\cos\theta_i \right) - \frac{\partial E_{\hat{z}}}{\partial \hat{y}} + \frac{\partial E_{\hat{y}}}{\partial \hat{z}}$$

$$\frac{\partial E_{\hat{y}}}{\partial \tau} = i\beta\left(E_{\hat{y}} + H_{\hat{x}}\cos\theta_i \right) + \frac{\partial H_{\hat{x}}}{\partial \hat{z}} - 4\pi(J_{\hat{y}} - i\beta P_{\hat{y}})$$

$$\frac{\partial E_{\hat{z}}}{\partial \tau} = i\beta\left(E_{\hat{z}} + H_{\hat{x}}\sin\theta_i \right) - \frac{\partial H_{\hat{x}}}{\partial \hat{y}} - 4\pi(J_{\hat{z}} - i\beta P_{\hat{z}})$$

$$\frac{\partial J_{\hat{y}}}{\partial \tau} = \left(2i\beta - \tilde{\gamma} \right) J_{\hat{y}} + \left(\beta^2 + i\tilde{\gamma}\beta - \beta_0^2 \right) P_{\hat{y}} + \frac{\pi\omega_p^2}{\omega_r^2} E_{\hat{y}} + P_{\hat{y}}^{NL} \qquad . \qquad (16)$$

$$\frac{\partial J_{\hat{z}}}{\partial \tau} = \left(2i\beta - \tilde{\gamma} \right) J_{\hat{z}} + \left(\beta^2 + i\tilde{\gamma}\beta - \beta_0^2 \right) P_{\hat{z}} + \frac{\pi\omega_p^2}{\omega_r^2} E_{\hat{z}} + P_{\hat{z}}^{NL}$$

$$\frac{\partial P_{\hat{y}}}{\partial \tau} = J_{\hat{y}}$$

$$\frac{\partial P_{\hat{z}}}{\partial \tau} = J_{\hat{z}}$$

We have used the scaled coordinates: $\hat{z} = z/\lambda_r$, $\hat{y} = y/\lambda_r$, $\hat{x} = x/\lambda_r$, $\tau = ct/\lambda_r$, scaled frequencies $\beta = 2\pi\omega/\omega_r$ (for the FF field), $\beta_0 = 2\pi\omega_0/\omega_r$, and damping coefficient $\tilde{\gamma} = \gamma(\lambda_r/c)$, where $\lambda_r = 1\mu m$ is a conveniently chosen reference wavelength such that

11

$\omega_r = 2\pi c / \lambda_r$; θ_i is the angle of incidence; $P_{\hat{y}}^{NL}$ and $P_{\hat{z}}^{NL}$ are generic, nonlinear polarization components yet to be determined. Eqs.(16) are solved in the time domain using a modified FFT pulse propagation method [23], which we discuss in some detail next. Examination of the electric and magnetic field equations shows that they may be put into a Schrödinger-like form, namely:

$$\frac{\partial H_{\hat{x}}}{\partial \tau} = V_{H_{\hat{x}}} H_{\hat{x}} - \frac{\partial E_{\hat{z}}}{\partial \hat{y}} + \frac{\partial E_{\hat{y}}}{\partial \hat{z}}$$

$$\frac{\partial E_{\hat{y}}}{\partial \tau} = V_{E_{\hat{y}}} E_{\hat{y}} + \frac{\partial H_{\hat{x}}}{\partial \hat{z}} \qquad . \qquad (17)$$

$$\frac{\partial E_{\hat{z}}}{\partial \tau} = V_{E_{\hat{z}}} E_{\hat{z}} - \frac{\partial H_{\hat{x}}}{\partial \hat{y}}$$

The potentials in Eqs.(17) are $V_{H_{\hat{x}}} = i\beta\left(1 + E_{\hat{z}} \sin\theta_i + E_{\hat{y}} \cos\theta_i\right)/H_{\hat{x}}$,

$V_{E_{\hat{y}}} = i\beta\left(1 + H_{\hat{x}} \cos\theta_i - 4\pi(J_{\hat{y}} - i\beta P_{\hat{y}})\right)/E_{\hat{y}}$, and $V_{E_{\hat{z}}} = i\beta\left(1 + H_{\hat{x}} \sin\theta_i - 4\pi(J_{\hat{z}} - i\beta P_{\hat{z}})\right)/E_{\hat{z}}$. As

a result, the equations are solvable using the classic, split-step, beam propagation method adapted for the time domain [23]. The split-step algorithm usually calls for separation of free-space and material equations with differential equations that are first order in time. Eqs.(17) are already first order in time, have no approximations, can be immediately separated into free-space and material equations, and integrated in the time domain. The formal solutions of the free-space propagator may be derived from the free-space equations, obtained by setting the effective potentials equal to zero. Then, Eqs.(17) are Fourier transformed in space resulting in:

$$\frac{\partial \tilde{H}_{\hat{x}}}{\partial \tau} = -ik_{\hat{y}}\tilde{E}_{\hat{z}} + ik_{\hat{z}}\tilde{E}_{\hat{y}}$$

$$\frac{\partial \tilde{E}_{\hat{y}}}{\partial \tau} = ik_{\hat{z}}\tilde{H}_{\hat{x}} \qquad . \qquad (18)$$

$$\frac{\partial \tilde{E}_{\hat{z}}}{\partial \tau} = -ik_{\hat{y}}\tilde{H}_{\hat{x}}$$

Eqs.(18) may be integrated simultaneously using a midpoint trapezoidal method, so that:

12

$$\tilde{H}_{\hat{x}}(\delta\tau) = \tilde{H}_{\hat{x}}(0) - \frac{ik_{\hat{y}}\delta\tau}{2}\left(\tilde{E}_{\hat{z}}(0) + \tilde{E}_{\hat{z}}(\delta\tau)\right) + \frac{ik_{\hat{z}}\delta\tau}{2}\left(\tilde{E}_{\hat{y}}(0) + \tilde{E}_{\hat{y}}(\delta\tau)\right)$$

$$\tilde{E}_{\hat{y}}(\delta\tau) = \tilde{E}_{\hat{y}}(0) + \frac{ik_{\hat{z}}\delta\tau}{2}\left(\tilde{H}_{\hat{x}}(0) + \tilde{H}_{\hat{x}}(\delta\tau)\right) \qquad (19)$$

$$\tilde{E}_{\hat{z}}(\delta\tau) = \tilde{E}_{\hat{z}}(0) - \frac{ik_{\hat{y}}\delta\tau}{2}\left(\tilde{H}_{\hat{x}}(0) + \tilde{H}_{\hat{x}}(\delta\tau)\right)$$

Solving for $\tilde{H}_{\hat{x}}(\delta\tau)$:

$$\tilde{H}_{\hat{x}}(\delta\tau) = \tilde{H}_{\hat{x}}(0)\frac{\left(1 - \frac{\left(k_{\hat{y}}^2 + k_{\hat{z}}^2\right)\delta\tau^2}{4}\right)}{\left(1 + \frac{\left(k_{\hat{y}}^2 + k_{\hat{z}}^2\right)\delta\tau^2}{4}\right)} + \frac{\left(ik_{\hat{z}}\tilde{E}_{\hat{y}}(0) - ik_{\hat{y}}\tilde{E}_{\hat{z}}(0)\right)\delta\tau}{\left(1 + \frac{\left(k_{\hat{y}}^2 + k_{\hat{z}}^2\right)\delta\tau^2}{4}\right)} , \qquad (20)$$

Eq.(20) is then substituted back into the second and third of Eqs.(19) to extract the electric fields. All fields are then inverse Fourier transformed.

The propagation step inside the medium is performed by integrating the material equations, also derived from Eqs.(17) and written in terms of generic envelope functions as:

$$\frac{\partial H_{\hat{x}}}{\partial\tau} = i\beta\left(H_{\hat{x}} + E_{\hat{z}}\sin\theta_i + E_{\hat{y}}\cos\theta_i\right)$$

$$\frac{\partial E_{\hat{y}}}{\partial\tau} = i\beta\left(E_{\hat{y}} + H_{\hat{x}}\cos\theta_i\right) - 4\pi(J_{\hat{y}} - i\beta P_{\hat{y}})$$

$$\frac{\partial E_{\hat{z}}}{\partial\tau} = i\beta\left(E_{\hat{z}} + H_{\hat{x}}\sin\theta_i\right) - 4\pi(J_{\hat{z}} - i\beta P_{\hat{z}})$$

$$\frac{\partial J_{\hat{y}}}{\partial\tau} = \left(2i\beta - \tilde{\gamma}\right)J_{\hat{y}} + \left(\beta^2 + i\tilde{\gamma}\beta - \beta_0^2\right)P_{\hat{y}} + \frac{\pi\omega_p^2}{\omega_r^2}E_{\hat{y}} + P_{\hat{y}}^{NL} . \qquad (21)$$

$$\frac{\partial J_{\hat{z}}}{\partial\tau} = \left(2i\beta - \tilde{\gamma}\right)J_{\hat{z}} + \left(\beta^2 + i\tilde{\gamma}\beta - \beta_0^2\right)P_{\hat{z}} + \frac{\pi\omega_p^2}{\omega_r^2}E_{\hat{z}} + P_{\hat{z}}^{NL}$$

$$\frac{\partial P_{\hat{y}}}{\partial\tau} = J_{\hat{y}}$$

$$\frac{\partial P_{\hat{z}}}{\partial\tau} = J_{\hat{z}}$$

Although we have neglected magnetic currents and polarizations, which typically characterize magnetically active and negative index materials, for example, they may be reintroduced in straightforward fashion by adding magnetic sources. Then, an approach similar to the solution of

13

Eq.(18) may be employed to solve Eqs.(21). For instance, one may first obtain estimates (the prediction step) of all fields, currents, and polarizations at $\tau = \delta\tau$ with an Euler method, using only their initial values at $\tau=0$. Using these predictions, the solutions for the currents at $\tau = \delta\tau$ are immediate and second-order accurate, as follows:

$$
J_{\hat{y},\hat{z}}(\delta\tau) = J_{\hat{y},\hat{z}}(0) \frac{\left(1 + (2i\beta - \tilde{\gamma})\dfrac{\delta\tau}{2} + (\beta^2 + i\tilde{\gamma}\beta - \beta_0^2)\dfrac{\delta\tau^2}{4}\right)}{\left(1 - (2i\beta - \tilde{\gamma})\dfrac{\delta\tau}{2} - (\beta^2 + i\tilde{\gamma}\beta - \beta_0^2)\dfrac{\delta\tau^2}{4}\right)}
$$
$$
+ \frac{(\beta^2 + i\tilde{\gamma}\beta - \beta_0^2)P_{\hat{y},\hat{z}}(0)\delta\tau + \pi\dfrac{\omega_P^2}{\omega_r^2}\left(E_{\hat{y},\hat{z}}(0) + E_{P,\hat{y},\hat{z}}(\delta\tau)\right)\dfrac{\delta\tau}{2} + \left(P_{\hat{y},\hat{z}}^{NL}(0) + P_{P,\hat{y},\hat{z}}^{NL}(\delta\tau)\right)\dfrac{\delta\tau}{2}}{\left(1 - (2i\beta - \tilde{\gamma})\dfrac{\delta\tau}{2} - (\beta^2 + i\tilde{\gamma}\beta - \beta_0^2)\dfrac{\delta\tau^2}{4}\right)} \quad . \quad (22)
$$

$E_{P,\hat{y},\hat{z}}(\delta\tau)$ and $P_{P,\hat{y},\hat{z}}^{NL}(\delta\tau)$ are first-order accurate, predicted estimates of the fields and nonlinear polarizations at time $\tau=\delta\tau$. Once the currents are known, the polarizations may be found using the usual, second-order-accurate trapezoidal rule:

$$
P_{\hat{y},\hat{z}}(\delta\tau) = P_{\hat{y},\hat{z}}(0) + \left(J_{\hat{y},\hat{z}}(\delta\tau) + J_{\hat{y},\hat{z}}(0)\right)\frac{\delta\tau}{2} \quad . \quad (23)
$$

In turn, knowledge of more accurate currents and polarizations at time $\delta\tau$ allows second-order accurate estimates of all electric and magnetic fields. The process is then repeated several times, although one or two cycles usually suffice for the results to converge (the correction step).

In order to solve the full nonlinear problem, the linear data represented by Eq.(12) may be reproduced dynamically in a numerical integration scheme by solving a set of coupled equations that contains: (1) an equation that describes the polarization of conduction electrons, $\mathbf{P}_f$; and (2) for this case two equations, each similar to Eq.(1), to account for the bound electron polarizations, $\mathbf{P}_1$, and $\mathbf{P}_2$, respectively. Bound-electron contributions are pivotal in the region

14

labeled "Lorentz" on Fig.1a [24, 25], both for linear and nonlinear contributions to the dielectric constant, and thus harmonic generation. The linear equations are then modified to include nonlinear contributions in the high intensity regime [24, 25]:

$$\ddot{\mathbf{P}}_f + \tilde{\gamma}_f \dot{\mathbf{P}}_f = \frac{n_{0f}e^2}{m_0^*}\left(\frac{\lambda_r}{c}\right)^2 \mathbf{E} + \frac{5}{3}\frac{E_F}{m_0^*c^2}\nabla\left(\nabla\bullet\mathbf{P}_f\right) - \frac{e\,\lambda_r}{m_0^*c^2}\mathbf{E}\left(\nabla\bullet\mathbf{P}_f\right) + \frac{e\,\lambda_r}{m_0^*c^2}\dot{\mathbf{P}}_f\times\mathbf{H}$$
$$-\frac{1}{n_{0f}e\lambda_r}\left[\left(\nabla\bullet\dot{\mathbf{P}}_f\right)\dot{\mathbf{P}}_f + \left(\dot{\mathbf{P}}_f\bullet\nabla\right)\dot{\mathbf{P}}_f\right] - \frac{10}{9}\frac{E_F}{m_0^*c^2}\frac{1}{n_{0f}e\lambda_r}\left(\nabla\bullet\mathbf{P}_f\right)\nabla\left(\nabla\bullet\mathbf{P}_f\right) \quad , \tag{24}$$

$$\ddot{\mathbf{P}}_1 + \tilde{\gamma}_{01}\dot{\mathbf{P}}_1 + \tilde{\omega}_{01}^2\mathbf{P}_1 - \tilde{b}_1\left(\mathbf{P}_1\bullet\mathbf{P}_1\right)\mathbf{P}_1 = \frac{n_{01}e^2\lambda_r^2}{m_{b1}^*c^2}\mathbf{E} + \frac{e\,\lambda_r}{m_{b1}^*c^2}\dot{\mathbf{P}}_1\times\mathbf{H} \ , \tag{25}$$

$$\ddot{\mathbf{P}}_2 + \tilde{\gamma}_{02}\dot{\mathbf{P}}_2 + \tilde{\omega}_{02}^2\mathbf{P}_2 - \tilde{b}_2\left(\mathbf{P}_2\bullet\mathbf{P}_2\right)\mathbf{P}_2 = \frac{n_{02}e^2\lambda_r^2}{m_{b2}^*c^2}\mathbf{E} + \frac{e\,\lambda_r}{m_{b2}^*c^2}\dot{\mathbf{P}}_2\times\mathbf{H} \quad . \tag{26}$$

To summarize, Eq.(24) amounts to an augmented Drude or hydrodynamic model that describes a nearly-free electron gas inside the metal; Eqs.(25-26) account for two separate bound electron species, each subject to internal and external forces. Each electron species has its own damping rate ($\tilde{\gamma}_f$, $\tilde{\gamma}_{01}$, $\tilde{\gamma}_{02}$), effective mass (m_0^*, m_{b1}^*, m_{b2}^*), density (n_{0f}, n_{01}, n_{02}), and in the case of bound electrons, resonance frequency ($\tilde{\omega}_{01}$, $\tilde{\omega}_{02}$), and associated, scaled, nonlinear third-order coefficient ($\tilde{b}_1 = b_1\dfrac{\lambda_r^2}{n_{01}^2e^2c^2}$, $\tilde{b}_2 = b_2\dfrac{\lambda_r^2}{n_{02}^2e^2c^2}$). For simplicity we choose to operate in two spatial dimensions plus time, where $\nabla \equiv \dfrac{\partial}{\partial\hat{z}}\hat{\mathbf{k}} + \dfrac{\partial}{\partial\hat{y}}\hat{\mathbf{j}}$, as the fields may be polarized in that direction but are assumed to be independent of the scaled coordinate $\hat{x}$. In addition to the usual linear driving term that characterizes a pure Drude model, i.e. the first term on the right hand side of Eq.(24), $\dfrac{n_{0f}e^2}{m_0^*}\left(\dfrac{\lambda_r}{c}\right)^2\mathbf{E}$, there are other free-electron contributions all leading to or influencing SHG and

15

THG, that may be summarized as follows: the nonlinear magnetic Lorentz force, $\dfrac{e\,\lambda_r}{m_0^*c^2}\dot{\mathbf{P}}_f\times\mathbf{H}$; an

explicit quadrupole-like, nonlinear Coulomb term that arises from the continuity

equation, $-\dfrac{e\,\lambda_r}{m_0^*c^2}\mathbf{E}\left(\nabla\bullet\mathbf{P}_f\right)$; two nonlinear convective terms,

$-\dfrac{1}{n_{0f}e\lambda_r}\left[\left(\nabla\bullet\dot{\mathbf{P}}_f\right)\dot{\mathbf{P}}_f+\left(\dot{\mathbf{P}}_f\bullet\nabla\right)\dot{\mathbf{P}}_f\right]$; a linear electron gas pressure (nonlocal) pressure term,

$\nabla\left(\nabla\bullet\mathbf{P}_f\right)$, and finally a nonlinear electron gas pressure terms, $\left(\nabla\bullet\mathbf{P}_f\right)\nabla\left(\nabla\bullet\mathbf{P}_f\right)$, having

strictly quantum origins. By the same token, the fields in Eqs.(25-26) may each be expanded

around an equilibrium, electron position, giving rise to additional, quadrupole-like, nonlinear

force terms arising from Coulomb sources [2, 24, 25]. Upon expansion into their harmonic

components, each of Eqs.(24-26) splits into three, separate equations. For instance, Eq.(25) may

be written as follows [24, 25]:

$$\ddot{\mathbf{P}}_{1,\omega}+\tilde{\gamma}_{1,\omega}\dot{\mathbf{P}}_{1,\omega}+\tilde{\omega}_{01,\omega}^2\mathbf{P}_{1,\omega}-\mathbf{P}_{1,\omega}^{(NL)}\approx\frac{n_{01}e^2\lambda_0^2}{m_{b1}^*c^2}\mathbf{E}_\omega+\frac{e\,\lambda_0}{m_{b1}^*c^2}\left(\begin{pmatrix}-\dfrac{1}{2}\mathbf{E}_\omega^*\nabla\bullet\mathbf{P}_{1,2\omega}\\[4pt]+2\mathbf{E}_{2\omega}\nabla\bullet\mathbf{P}_{1,\omega}^*\\[4pt]-\dfrac{2}{3}\mathbf{E}_{2\omega}^*\nabla\bullet\mathbf{P}_{1,3\omega}\\[4pt]-\dfrac{3}{2}\mathbf{E}_{3\omega}\nabla\bullet\mathbf{P}_{1,2\omega}^*\end{pmatrix}+\begin{pmatrix}\left(\dot{\mathbf{P}}_{1,\omega}^*+i\omega\mathbf{P}_{1,\omega}^*\right)\times\mathbf{H}_{2\omega}\\[4pt]+\left(\dot{\mathbf{P}}_{1,2\omega}-2i\omega\mathbf{P}_{1,2\omega}\right)\times\mathbf{H}_\omega^*\\[4pt]+\left(\dot{\mathbf{P}}_{1,2\omega}^*+2i\omega\mathbf{P}_{1,2\omega}^*\right)\times\mathbf{H}_{3\omega}\\[4pt]+\left(\dot{\mathbf{P}}_{1,3\omega}-3i\omega\mathbf{P}_{1,3\omega}\right)\times\mathbf{H}_{2\omega}^*\end{pmatrix}\right)$$

$$\ddot{\mathbf{P}}_{1,2\omega}+\tilde{\gamma}_{1,2\omega}\dot{\mathbf{P}}_{1,2\omega}+\tilde{\omega}_{02,2\omega}^2\mathbf{P}_{1,2\omega}-\mathbf{P}_{1,2\omega}^{(NL)}\approx\frac{n_{01}e^2\lambda_0^2}{m_{b1}^*c^2}\mathbf{E}_{2\omega}+\frac{e\,\lambda_0}{m_{b1}^*c^2}\left(\begin{pmatrix}\mathbf{E}_\omega\nabla\bullet\mathbf{P}_{1,\omega}\\[4pt]-\dfrac{1}{3}\mathbf{E}_\omega^*\nabla\bullet\mathbf{P}_{1,3\omega}\\[4pt]-3\mathbf{E}_{3\omega}\nabla\bullet\mathbf{P}_{1,\omega}^*\end{pmatrix}+\begin{pmatrix}\left(\dot{\mathbf{P}}_{1,\omega}-i\omega\mathbf{P}_{1,\omega}\right)\times\mathbf{H}_\omega\\[4pt]+\left(\dot{\mathbf{P}}_{1,\omega}^*+i\omega\mathbf{P}_{1,\omega}^*\right)\times\mathbf{H}_{3\omega}\\[4pt]+\left(\dot{\mathbf{P}}_{1,3\omega}-3i\omega\mathbf{P}_{1,3\omega}\right)\times\mathbf{H}_\omega^*\end{pmatrix}\right)$$

$$\ddot{\mathbf{P}}_{1,3\omega}+\tilde{\gamma}_{1,3\omega}\dot{\mathbf{P}}_{1,3\omega}+\tilde{\omega}_{03,3\omega}^2\mathbf{P}_{1,3\omega}-\mathbf{P}_{1,3\omega}^{(NL)}\approx\frac{n_{01}e^2\lambda_0^2}{m_{b1}^*c^2}\mathbf{E}_{3\omega}+\frac{e\,\lambda_0}{m_{b1}^*c^2}\left(\begin{pmatrix}\dfrac{1}{2}\mathbf{E}_\omega\nabla\bullet\mathbf{P}_{1,2\omega}\\[4pt]+2\mathbf{E}_{2\omega}\nabla\bullet\mathbf{P}_{1,\omega}\end{pmatrix}+\begin{pmatrix}\left(\dot{\mathbf{P}}_{1,2\omega}-2i\omega\mathbf{P}_{1,2\omega}\right)\times\mathbf{H}_\omega\\[4pt]+\left(\dot{\mathbf{P}}_{1,\omega}-i\omega\mathbf{P}_{1,\omega}\right)\times\mathbf{H}_{2\omega}\end{pmatrix}\right)$$

, (27)

and similarly for Eq.(26). Finally, having opted for a dynamic description of the nonlinear response with the nonlinear term $\mathbf{P}_{1,2}^{(NL)} = \tilde{b}_{1,2}\left(\mathbf{P}_{1,2} \bullet \mathbf{P}_{1,2}\right)\mathbf{P}_{1,2}$ in order to retain simultaneously linear and nonlinear dispersions, the general expansion of the third-order nonlinear polarization for just *one of the bound oscillators* becomes:

$$P_{i,1}^{NL} = \sum_{j=1,3} \sum_{k=1,3} \sum_{l=1,3} \tilde{b}_{ijkl,1} P_{j,1} P_{k,1} P_{l,1} \qquad , \qquad (28)$$

and similarly for the second oscillator. $\tilde{b}_{ijkl,1}$ is now a tensor, in full analogy with the expansion of the nonlinear polarizations in terms of χ_{ijkl} [1]. For example, for a material like GaAs, which has cubic symmetry of the type $\overline{4}3m$, Eq.(28) reduces to:

$$
\begin{aligned}
P_{x,1}^{NL} &= \tilde{b}_{xxxx,1}^{(3)} P_{x,1}^3 + 3\tilde{b}_{xxyy,1}^{(3)} P_{y,1}^2 P_{x,1} + 3\tilde{b}_{xxzz,1}^{(3)} P_{z,1}^2 P_{x,1} \\
P_{y,1}^{NL} &= \tilde{b}_{yyyy,1}^{(3)} P_y^3 + 3\tilde{b}_{xxyy,1}^{(3)} P_{x,1}^2 P_{y,1} + 3\tilde{b}_{yyzz,1}^{(3)} P_{z,1}^2 P_{y,1} \qquad (29) \\
P_{z,1}^{NL} &= \tilde{b}_{zzzz,1}^{(3)} P_{z,1}^3 + 3\tilde{b}_{zzxx,1}^{(3)} P_{x,1}^2 P_{z,1} + 3\tilde{b}_{zzyy,1}^{(3)} P_{y,1}^2 P_{z,1}
\end{aligned}
$$

The situation is comparable for noble metals, except that for isotropic crystal symmetry the relations between the tensor components allow one to more simply write:

$$
\begin{aligned}
P_{x,1}^{NL} &= \tilde{b}_1^{(3)}\left(P_{x,1}^3 + P_{y,1}^2 P_{x,1} + P_{z,1}^2 P_{x,1}\right) \\
P_{y,1}^{NL} &= \tilde{b}_1^{(3)}\left(P_{y,1}^3 + P_{x,1}^2 P_{y,1} + P_{z,1}^2 P_{y,1}\right) \qquad . \qquad (30) \\
P_{z,1}^{NL} &= \tilde{b}_1^{(3)}\left(P_{z,1}^3 + P_{x,1}^2 P_{z,1} + P_{y,1}^2 P_{z,1}\right)
\end{aligned}
$$

Considerable care should be exercised when decomposing Eqs.(30) into its harmonics, and while substituting those components (which include up- and down-conversion terms, self- and cross-phase modulation terms) back into Eqs.(24-27), thus making these equations valid even under conditions of pump depletion.

3. Results

In Fig. 2 we show the structures we use for the application of our model: a bare silver layer (top panel), a symmetric, transparent metal-dielectric stack (middle panel) [26, 27], and an opaque plasmonic grating (bottom panel). No effort is made to optimize conversion efficiency of one structure with respect to another. Rather, these structures are chosen primarily for their simplicity, because in at least two cases, the thin metal layer and the multilayer stack, no experimental results have yet been reported, and also because the layered structures offer something that plasmonic structures usually do not, namely large field penetration depth inside the bulk of the metal. In contrast, plasmonic structures present us with large local fields near the surface of the metal, leaving the evanescent tail of the field to harvest the effects of the bulk's nonlinearity. These choices thus alleviate the burden on an eventual experimental effort and should make a subsequent comparison with predictions simpler to interpret. Each panel contains the corresponding linear transmission, reflection, and absorption spectra obtained at normal incidence. Worthy of note are: (i) the typical broadband transmission spectrum of the multilayer stack, which may be extended well into the infrared range; and (ii) the grating's narrow plasmonic reflection resonance. The stack's high transmittance ensures that the electric fields propagate and localize inside each metal layer, thus creating an opportunity to exploit the metal's bulk nonlinearity [28, 29, 30]. By the same token, at resonance the plasmonic grating confines the field in close
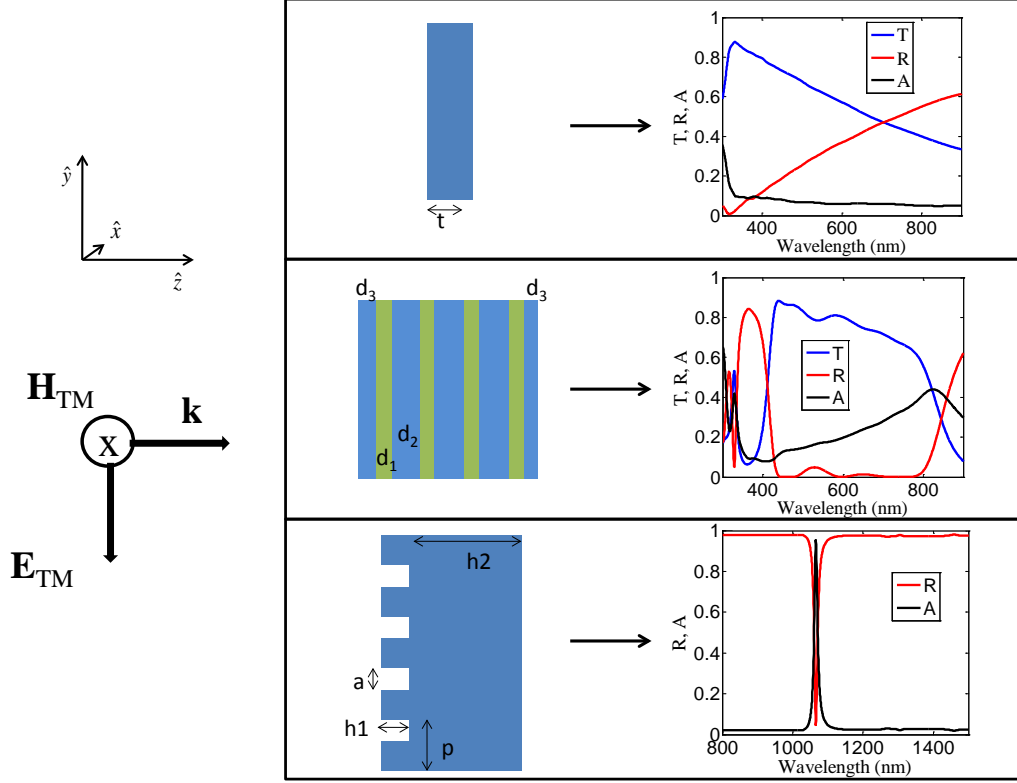
Fig.2: Top panel: a single silver layer, t=10nm; Middle panel: symmetric metal-dielectric stack composed of four Ag layers of thickness d_1=12.5nm, and Ta_2O_5 layers such that d_2=80nm and d_3=40nm; Bottom panel: Silver grating with p=1030nm, a=300nm, h_1=45nm, h_2=200nm. The respective transmission, reflection, and absorption spectra at normal incidence are depicted to the right of each object. The field is incident from the left, and polarized as shown.

proximity of the surface, producing an intense field near internal and external corners. A bound wave travels along the grating, with an evanescent tail that spills just inside the metal, which in turn absorbs all the incident energy resulting in near-zero reflections. In Fig. 3 we show the results of vectorial, nonlinear calculations performed using incident Gaussian pulses approximately 50fs in duration for the layered structures, and approximately 700fs for the grating, in order to resolve the narrow resonance. Peak intensities are ~2GW/cm^2,

and $\tilde{b}_1 = \tilde{b}_2 = \dfrac{\omega_0^2 \lambda_r^2}{r_0^2 n_{01}^2 e^2 c^2} = 10^{-6}$. We compare reflected, TH conversion efficiencies at normal
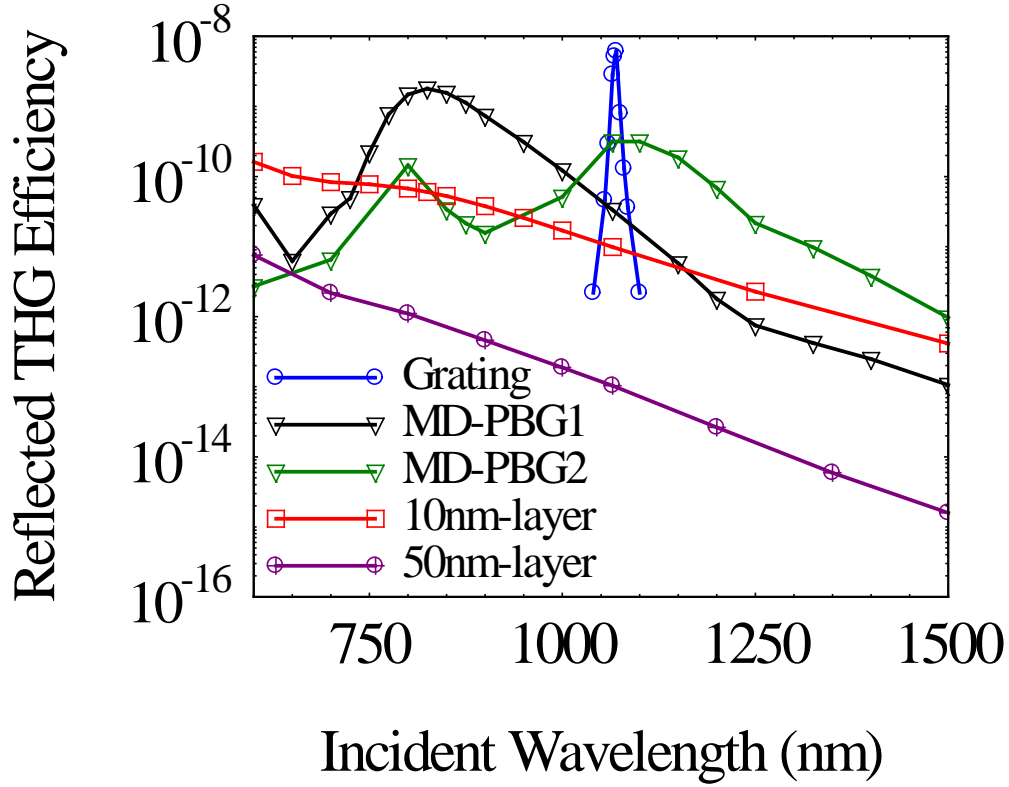
Fig.3: Reflected THG conversion efficiency vs. incident pump wavelength for the three structures depicted in Fig.2. The calculations are carried out for two free-standing silver layers to highlight the dramatic impact of thickness on harmonic generation, and for two slightly different multilayer stacks. The multilayer stack labeled MD-PBG1 is the one illustrated in Fig.2. The curve labeled MD-PBG2 refers to a stack with silver layers are $d_1=12.5$nm, and Ta_2O_5 layers are $d_2=125$nm and $d_3=62.5$nm. Thicker Ta_2O_5 layers red-shift the band gap structure to yield maximum absorption near 1070nm, for a fairer comparison with the grating. All things being equal, the plasmonic grating outperforms the multilayer stack by a mere order of magnitude. The stack compensates the larger, plasmonic-assisted field enhancement by penetrating inside the metal layers to take full advantage of its large nonlinearity.

incidence for the three types of structures depicted in Fig. 2. The increased conversion efficiency as a function of decreasing wavelength for the smooth 10nm- and 50nm-tick metal layers may be understood simply in terms of increased penetration depth, which is more pronounced for thinner layers and shorter wavelengths, or a combination of both. However, in the visible range two orders of magnitude separate the predicted TH conversion efficiencies of the two metal layers, while the discrepancy is three orders of magnitudes at 1500nm, increasing further for thicker layer. The dependence of conversion efficiency on layer thickness may be ascertained from Fig. 4 for a representative incident wavelength of 850nm.
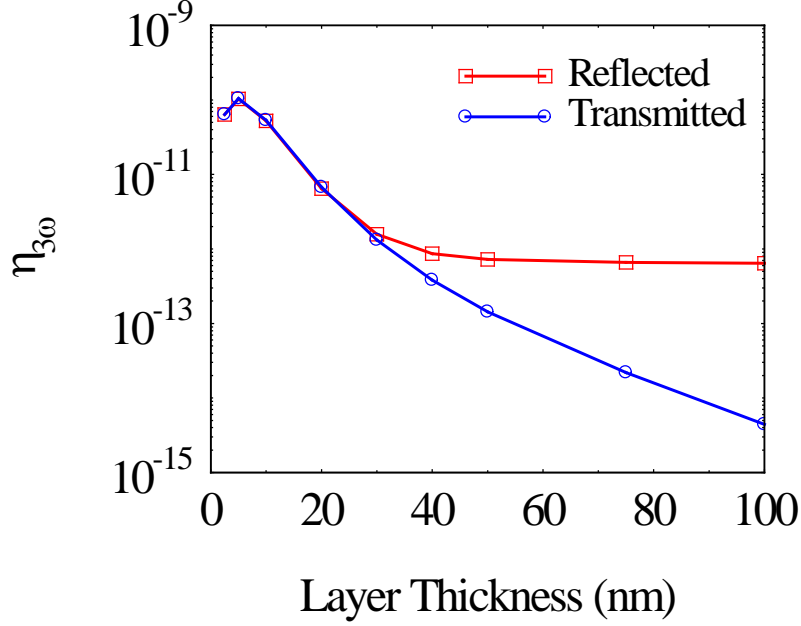
Fig.4: Reflected and transmitted THG conversion efficiency vs. silver layer thickness, assuming the pump is tuned at 850nm; the layer is assumed to remain uniform and the dielectric constant retains the bulk characteristics reported in Palik's handbook, even for very thin layers. At this wavelength, maximum THG occurs for layer thickness of ~5nm, reflected THG saturates at ~50nm, and transmitted THG decreases exponentially with thickness.

A comparison of the medal-dielectric stack's linear spectra in Fig. 2 with the curve labeled MD-PBG1 in Fig. 3 (i.e., its THG conversion efficiency) reveals that the linear absorption peak near 850nm coincides with an absolute maximum in THG. The situation repeats for the curve labeled MD-PBG2, whose geometry is similar to that of MD-PBG1, and is described in the caption of Fig. 3, and again for the grating. This behavior is not unique to these systems, and it simply highlights the sensitivity of both linear absorption ($\sim \text{Im}(\varepsilon(\omega))|\mathbf{E}_\omega|^2$) and nonlinear frequency conversion ($\sim \chi_{3\omega}^{(3)}\mathbf{E}_\omega^2$) to field intensity.

The results in Fig. 3 suggest that plasmon-assisted THG mediated by the grating yields the largest predicted conversion efficiencies among the structures we have examined. However, it is relatively narrow-band, and it exceeds both stacks' conversion efficiencies by a mere order of magnitude. The figure also suggests that off-resonance even the 10nm-thick metal layer out-

performs the plasmonic grating by several orders of magnitudes. When it comes to nonlinear frequency conversion in plasmonic structures, off-resonant conditions [31] and isolated nanoparticles [32] are often employed as typical benchmarks to highlight extraordinary enhancement or performance. However, Fig. 3 suggest that these choices often lead to scenarios that may give the impression of performance that is far more favorable than more realistic benchmarks actually suggest, such as either a 5nm- or a 10nm-thick silver layer.

The angular dependence of THG from thick silver and aluminum mirrors was predicted and experimentally reported for an incident wavelength of 1064nm in reference [33] using 4.5ps pulses that were part of a pulse train hundreds of nanoseconds in duration. Those predictions were based on knowledge of the linear dielectric response [10, 11] and the calculation of the Fresnel coefficients, with a reflected THG peak near 63° for TM-polarized light. Our calculations are in good qualitative and quantitative agreement with those in reference [33] for incident wavelength of 1064nm and both TE- and TM-polarized light. Results for λ_{in}=850nm are plotted in Fig. 5(a) for TE- and TM-polarizations. The peak of maximum TH conversion efficiency for incident TM-polarized light matches the linear absorption maximum that occurs as a result of increased field penetration with respect to TE-polarized light, and shifts gradually to smaller angles as incident pump wavelength decreases, reaching ~56° at 632nm.

The qualitative differences that we predict between thick and thin metal layers are apparent: the 10nm-thick layer − Fig. 5(b) − displays maximum conversion efficiency at normal incidence, once again matching the linear absorption data obtained from the Fresnel coefficients, which are now similar for TE-and TM-polarizations, along with a concurrent increase by two orders of magnitude in conversion efficiency compared to the 100nm-thick layer. The curves in Fig. (5c) show that reflected conversion efficiencies from the multilayer stack are approximately

three orders of magnitude larger compared to the 100nm-thick layer, and one order of magnitude larger compared to the 10nm-thick silver layer, thanks to simultaneously improved field penetration and larger local fields for both polarizations, that
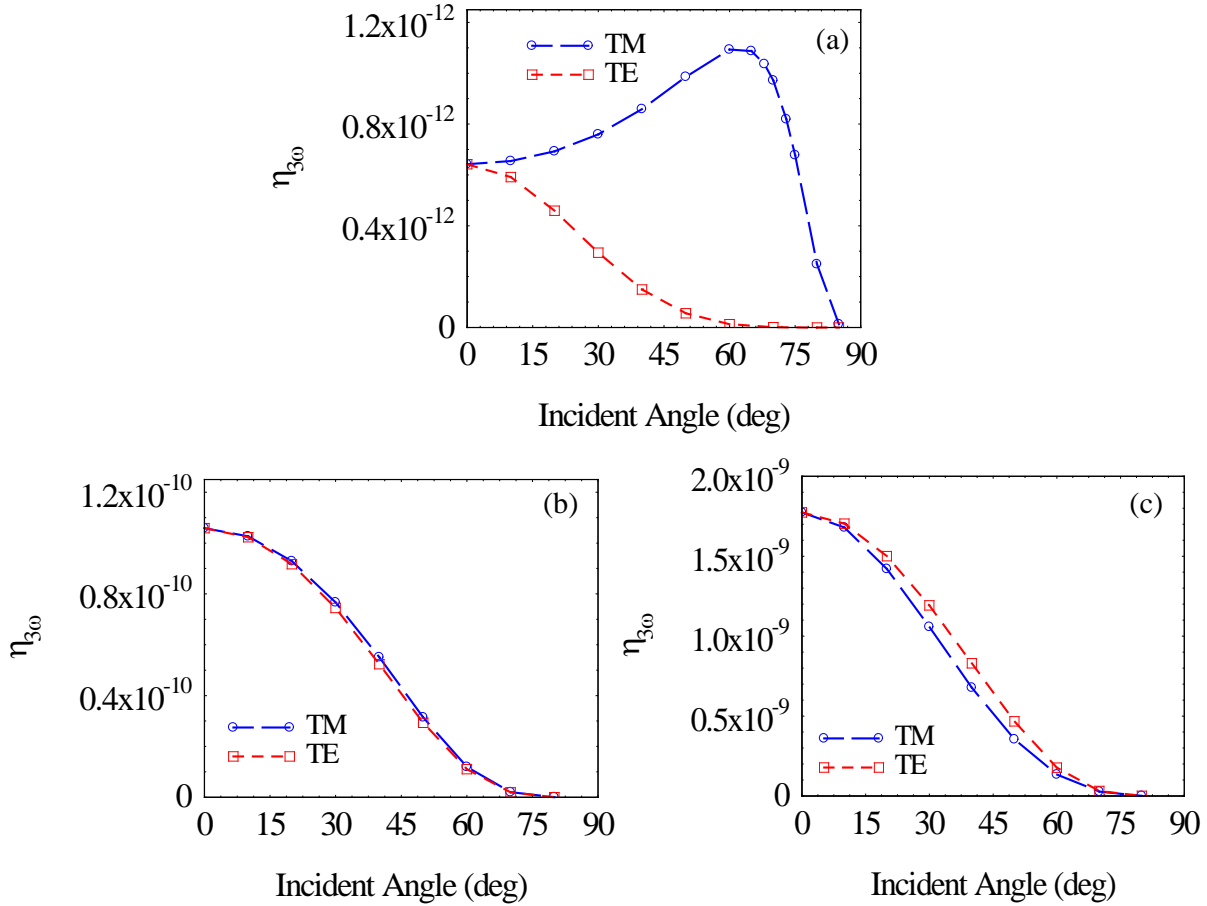

Fig.5: TM- and TE-polarized reflected THG conversion efficiency vs. incident angle for: (a) 100nm-thick and (b) 10nm-thick silver layers; (c) the metal-dielectric stack depicted in Fig.2. The pump is tuned at 850nm. Only the thick metal mirror yields discernable differences in conversion efficiencies between the two polarization states. THG transmission curves are not shown because transmission is almost identical to the THG reflection curve for the 10nm-thick layer (b), and negligible for both the 100nm-thick mirror (a) and the multilayer stack (c).

tend to concentrate inside the metal and thus to exploit its bulk nonlinearity. Finally, we adopt the nonlinear polarization model that assumes $\mathbf{P}^{NL} = \chi^{(3)}\mathbf{E}^3$ and compare conversion efficiencies obtained using the nonlinear oscillator model, Eq.(1). Our calculations show that the choice

23

$\tilde{b}_1 = \tilde{b}_2 = 10^{-6}$ corresponds to $\chi^{(3)} \approx 10^{-19} (\text{m/V})^2$ in the non-resonant region, as predicted by Eqs.(10-11), Figs. 2(c)-(d), if bound-electron effective masses are chosen as suggested above. This result is in relatively good agreement with the results obtained on silver mirrors in reference [33].

For completeness, in Fig. 6 we show the results for SHG for the structures we have investigated. Just as was the case for THG, the grating— Fig. 6(d) —, which is illuminated at normal incidence, yields the largest narrow-band conversion efficiency at the plasmonic resonance near 1070nm, and it is two orders of magnitude larger than the reflected SHG arising from the 100nm-thick silver mirror, Fig. 6(c). The silver mirror's maximum conversion efficiency occurs near 68°, and is larger than the conversion efficiency peaks generated by either the 10nm-thick layer— Fig. 6(a) —or the transparent multilayer stack — Fig. 6(b) —, notwithstanding improved field penetration, testimony to the fact that SHG remains mostly a surface phenomenon. The fact that maximum conversion efficiency of the arrangements that contains only flat metal layer occurs off-axis is due to the nature of the intrinsic second order nonlinearities [34, 35], is related to the maximum longitudinal field discontinuity, and brings up the question of what is a sensible benchmark for the purpose of comparing conversion efficiencies and assessing performance. A reasonable benchmark for SHG arising from the grating is probably the peak produced by the simplest, most efficient structure, which in this case appears to be the silver mirror peak at 68°. We should note that these considerations are not always made, resulting in rather suggestive claims [36]. For example, quantum-well multilayer stacks of the type discussed in reference [36] display a non-zero $\chi^{(2)}_{zzz}$ along the (longitudinal) direction of propagation. Under these conditions, it may be easily demonstrated that the bare

24

sample emits a SH signal most efficiently at large angles, and none at all at normal incidence, similarly to what happens in metals, because the effective second order nonlinearity arising from combined surface and volume sources has a dominant, longitudinal component. If the quantum well is then dressed with a metallic metasurface and illuminated at normal incidence, the more appropriate reference point for comparison should probably be the optimal conditions displayed by the bare sample, i.e. the SH peak that occurs at large incident angles, rather than what occurs at normal incidence. More realistic comparisons of the type we have discussed clearly suggest that photonic band gap structures, resonant plasmonic gratings and metasurfaces may indeed provide
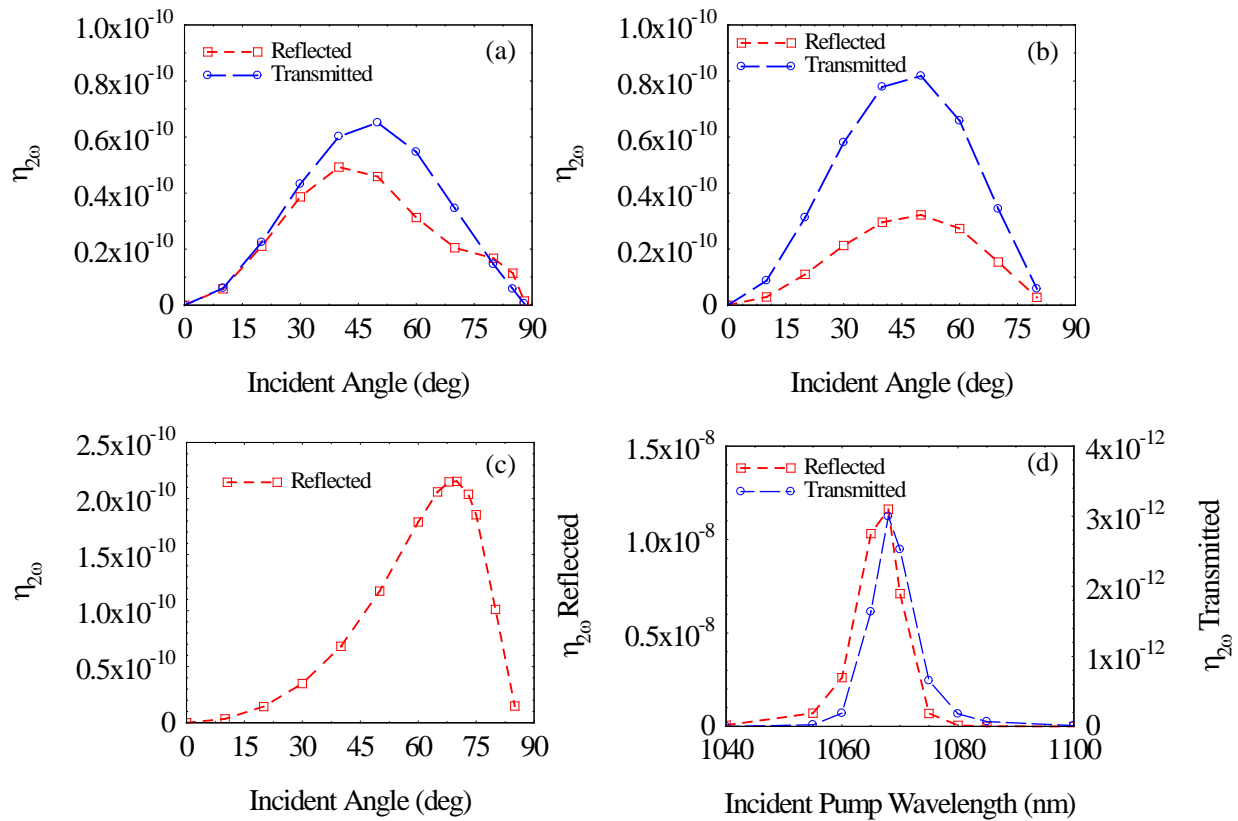

Fig.6: (a) SHG vs. incident angle for 10nm-thick silver; (b) the multilayer stack described in the caption of Fig.2; and (c) a 100nm-thick silver mirror. Both the 10nm layer and the multilayer stack display significant levels of transmitted SHG, while transmitted SHG for the 100nm-thick layer is of order 10^{-13}. The incident wavelength is 850nm in (a)-(c). (d) Transmitted and reflected SHG spectra for the grating described in Fig.2. In this case incident

pulse duration is approximately 700fs, in order to resolve the resonance. The overall enhancement the grating provides with respect to the flat-layered structure is nearly two orders of magnitude, which may be considered typical for either plasmonic or photonic band gap structures.
a few orders of magnitude of enhanced performance thanks to a combination of improved field enhancement and penetration depth, provided one compares to sensibly chosen benchmarks.

4. Conclusions

In conclusion, we have used a nonlinear Lorentz-Duffing oscillator model to predict second and third harmonic conversion efficiencies from isolated metal layers, metal-dielectric multilayer stacks, and a metal grating that do not display an intrinsic second-order bulk nonlinearity. One of the main advantages of using the nonlinear oscillator model is that it naturally accounts for linear and nonlinear dispersions simultaneously. At the same time we have also illustrated a numerical integration scheme that combines fast Fourier transforms and a modified predictor-corrector method to reliably and accurately integrate the relevant nonlinear, coupled Maxwell-Lorentz system without the need to introduce either artificial nonlinearities or saturable absorption. Our results suggest that the plasmonic metal gratings may indeed outperform all other structures investigated, albeit at the cost of bandwidth and increased sensitivity to fabrication imperfections. Increased penetration depth and field localization inside the metal layers boost the performance of layered structures across a broad range of wavelengths, suggesting that reasonable and appropriate benchmarks should be sought when comparing the enhancement of plasmonic gratings to other types of structures.

References

[1] R. W. Boyd, *"Nonlinear Optics,"* (Academic, 2003).

[2] E. Adler, "Nonlinear optical frequency polarization in a dielectric," Phys. Rev. **134,** A728 (1964).

[3] J. W. Haus, D. de Ceglia, M. A. Vincenti and M. Scalora, "Quantum Conductivity for Metal-Insulator-Metal Nanostructures," J. Opt. Soc. Am. B **31**, 259 (2014)

[4] J. W. Haus, D. de Ceglia, M. A. Vincenti and M. Scalora, "Nonlinear Quantum Tunneling Effects in Nano-Plasmonic Environments: Two-Photon Absorption and Harmonic Generation," J. Opt. Soc. Am. B **31**, A13-A19 (2014).

[5] M. Scalora, D. de Ceglia, M. A. Vincenti and J. W. Haus, "Nonlocal and Quantum Tunneling Contributions to Harmonic Generation in Nanostructures: Electron Cloud Screening Effects", Phys. Rev. A **90**, 013831 (2014).

[6] C. Ciracì, M. Scalora and D. R. Smith, "Third-harmonic generation in the presence of classical nonlocal effects in gap-plasmon nanostructures," Phys. Rev. B **90**, 205403 (2015).

[7] O. L. de Lange and R. E. Raab, "Surprises in the multipole description of macroscopic electrodynamics," Am. J. Phys. **74,** 301-312 (2006).

[8] J. D. Jackson, *"The Classical Electromagnetic field,"* (Wiley, 1999).

[9] J. A. Goldstone and E. Garmire, "Intrinsic Optical Bistability in Nonlinear media," Phys. Rev Lett. **53**, 910-913 (1984).

[10] R. C. Miller, "Optical Second harmonic Generation in Piezoelectric Crystals," Appl. Phys. Letters **5**, 17-18 (1964)

[11]   C. G. Garrett and F. N. H. Robinson, "Miller's Phenomenological Rule for Computing Nonlinear Susceptibilities," IEEE Jour. of Quantum Elect. **2**, 328-329 (1966).

[12]   D. Gordon, M. Helle and J. Peňano, "Fully explicit nonlinear optics model in a particle-in-cell framework," Comput. Phys. **250**, 388–402 (2013).

[13]   C. Varin, G. Bart, R. Emms and T. Brabec, "Saturable Lorentz model for fully explicit three-dimensional modeling of nonlinear optics," Opt. Express **23**, 2686-2695 (2015).

[14]   L. Allen and J. H. Eberly, "Optical Resonance and Two-Level Atoms," (Dover, 1987).

[15]   J. Koga, "Simulation Model for the Effects of Nonlinear Polarization on the Propagation of Intense Pulse Lasers," Opt. Letters **24**, 408-410 (1999).

[16]   C. Conti, A. Di Falco and G. Assanto, "Frequency Generation within the Forbidden Band Gap: All Optical Rabi-like Splitting in Photonic Crystals and Microcavities," Phys. Rev. E **70**, 066614 (2004).

[17]   M. Kaupp, "The Role of Radial Nodes of Atomic Orbitals for Chemical Bonding and the Periodic Table," Jour. of Comp. Chemistry **28**, 320-325 (2007).

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

[19]   A. M. A. Ibrahim and P. K. Choudury, "On the Maxwell-Duffing Approach to Model Photonic Deflection Sensor," IEEE Photon. Journal **5**, 6800812 (2013).

[20]   H. Ehrenreich and H. R. Philipp, "Optical properties of Ag and Cu", Phys. Rev. **128**, 1622 (1962).

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

[22]   R. W. Boyd, Z. Shi and I. De Leon, "The Third order Nonlinear Optical Susceptibility of Gold," Opt. Comm. 326, 74-79 (2014).

[23] M. Scalora and M. E. Crenshaw, "A beam propagation method that handles reflections", Opt. Comm. **108**, 191 (1994).

[24] M. Scalora, M. A. Vincenti, D. de Ceglia, V. Roppo, M. Centini, N. Akozbek and M. J. Bloemer, "Second- and third-harmonic generation in metal-based structures," Phys. Rev. A **82**, 043828 (2010).

[25] M. A. Vincenti, D. De Ceglia, V. Roppo and M. Scalora, "Harmonic generation in metallic, GaAs-filled nanocavities in the enhanced transmission regime at visible and UV wavelengths," Opt, Express 19, 2064-2078 (2011).

[26] M. Scalora, G. D'Aguanno, N. Mattiucci, M. J. Bloemer, D. de Ceglia, M. Centini, A. Mandatori, C. Sibilia, N. Akozbek, M. G. Cappeddu, M. Fowler and J. W. Haus, "Negative refraction and sub-wavelength focusing in the visible range using transparent metallo-dielectric stacks," Opt. Express **15**, 508-523 (2007).

[27] D. de Ceglia, M. Vincenti, M. Cappeddu, M. Centini, N. Akozbek, A. D'Orazio, J. Haus, M. Bloemer and M. Scalora, "Tailoring metallodielectric structures for superresolution and superguiding applications in the visible and near-ir ranges," Phys. Rev. A **77**, 033848 (2008).

[28] R. S. Bennink, Y.K. Yoon, R. W. Boyd and J. E. Sipe, "Accessing the optical nonlinearity of metals with metal–dielectric photonic bandgap structures," Opt. Lett. **24**, 1416 (1999).

[29] N. N. Lepeshkin, A. Schweinsberg, G. Piredda, R. S. Bennink and R. W. Boyd, "Enhanced Nonlinear Optical Response of One-Dimensional Metal-Dielectric Photonic Crystals," Phys. Rev. Lett. **93**, 123902 (2004).

[30]   M. Scalora, N. Mattiucci, G. D'Aguanno, M. C. Larciprete and M. J. Bloemer, "Nonlinear pulse propagation in one-dimensional metal-dielectric multilayer stacks: Ultrawide bandwidth optical limiting," Phys. Rev E **73**, 016603 (2006).

[31]   K. Li, X. Li, D. Y. Lei, S. Wu and Y. Zhan, "Plasmon Gap Mode-Assisted Third-Harmonic Generation from Metal Film-coupled Nanowires," Appl. Phys. Lett. **104**, 261105 (2014).

[32]   H. Aouani, M. Rahmani, M. Navarro-Cía and S. A. Maier, "Third-Harmonic-Upconversion Enhancement from a Single Semiconductor Nanoparticle Coupled to a Plasmonic Antenna," Nature Nanotechnology **9**, 290-294 (2014).

[33]   W. K. Burns and N. Bloembergen, "Third-harmonic generation in absorbing media of cubic or isotropic symmetry," Phys. Rev. B **4**, 3437-3450 (1971).

[34]   J. E. Sipe, V. C. Y. So, M. Fukui and G. I. Stegeman, "Analysis of second-harmonic generation at metal surfaces", Phys. Rev. B **21**, 4389 (1980).

[35]   D. Krause, C. W. Teplin, and C. T. Rogers, "Optical surface second harmonic measurements of isotropic thin-film metals: Gold, silver, copper, aluminum, and tantalum", J. Appl. Phys. **96**, 3626 (2004).

[36]   J. Lee, M. Tymchenko, C. Argyropoulos, P. Y. Chen, F. Lu, F. Demmerle, G. Boehm, M. C. Amann, A. Alù and M. A. Belkin, "Giant Nonlinear Response from Plasmonic Metasurfaces Coupled to Intersubband Transitions," *Nature* **511**, 65-69 (2014).