Free electron nonlinearities in heavily doped semiconductors plasmons

Federico De Luca, 1, 2 Michele Ortolani, 3, 4 and Cristian Ciracì1–3

1 Istituto Italiano di Tecnologia, Center for Biomolecular Nanotechnologies, Via Barsanti 14, 73010 Arnesano, Italy
2 Dipartimento di Matematica e Fisica “E. De Giorgi”, Università del Salento, via Arnesano, 73100 Lecce, Italy
3 Department of Physics, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185 Rome, Italy
4 Istituto Italiano di Tecnologia, Center for Life NanoSciences, Viale Regina Elena 291, 00161 Rome, Italy
(Dated: December 14, 2020)

Heavily doped semiconductors have emerged as tunable low-loss plasmonic materials at mid-infrared frequencies. In this article we investigate nonlinear optical phenomena associated with high concentration of free electrons. We use a hydrodynamic description to study free electron dynamics in heavily doped semiconductors up to third-order terms, which are usually negligible for noble metals. We find that cascaded third-harmonic generation due to second-harmonic signals can be as strong as direct third-harmonic generation contributions even when the second-harmonic generation efficiency is zero. Moreover, we show that when coupled with plasmonic enhancement free electron nonlinearities could be up to two orders of magnitude larger than conventional semiconductor nonlinearities. Our study might open a new route for nonlinear optical integrated devices at mid-infrared frequencies.

I. INTRODUCTION

Plasmonic nanoantennas emerged as a disruptive technology for concentrating and controlling light at the nanoscale thanks to their unique capabilities of field localization and enhancement [16–18]. Nonetheless, their widespread application has struggled to concretize because of several shortcomings of the main constituent of plasmonic materials: noble metals. The use of noble metals for practical devices has been limited due to, above all, high losses, high reflectivity and very poor compatibility with standard microelectronics manufacturing processes. In this context, heavily doped semiconductors (i.e. with charge densities $n_0 \sim 10^{19}$–$10^{20}$ cm$^{-3}$) have been introduced as low-loss and tunable alternative materials for plasmonics in the near-infrared (NIR), i.e. $0.8 < \lambda < 2 \mu$m, and in the mid-infrared (MIR), i.e. $2 < \lambda < 20 \mu$m [7, 8], being $\lambda$ the free-space wavelength. The MIR overlaps with the so-called molecular fingerprint region, and is particularly important for sensing applications in many fields, from healthcare to security [9]. Moreover, interest in MIR devices is growing in the context of free-space communications, since the 3–5 $\mu$m and the 8–13 $\mu$m regions are less affected by atmospheric turbulence, clouds and fog than the NIR region [10]. Because semiconductors offer the possibility of tuning their optical response through doping or electrical and optical excitation, and some of them are also highly compatible with silicon CMOS technology, their use for plasmonics may lead to breakthrough technologies [9].

Plasmonic systems owe their properties to surface plasmon modes – the resonant collective oscillations of free electrons (FEs) – that appear in materials with a high carrier concentration (i.e. metals and heavily doped semiconductors) as a consequence of the interaction with an external light excitation. Naturally, the local field enhancement provided by these modes especially affects optical processes whose dependence on the field intensity is nonlinear. Indeed, this is extremely important for nonlinear optical effects since they are in general very weak and normally require high laser intensities and long propagation distances in macroscopic nonlinear crystals to reach operational efficiencies. Although, recently, high-index dielectric resonators have been used to partially overcome these limitations [11, 12], plasmonic nanoantennas made of noble metals have been commonly used as local field amplifier in hybrid systems to boost optical nonlinearity from dielectric material placed in their vicinity [13, 14]. The nonlinear response may also originate directly from the plasmonic material itself [28, 29] and, in particular, from the dynamics of non-equilibrium FEs [31, 32]. For instance, in noble metals, FE plasmonic nonlinearities have been shown to strongly contribute to second-order processes [31, 37, 38] in the visible and NIR, and experimental measurements in gold nanoparticle arrays have demonstrated second-harmonic generation (SHG) efficiencies comparable to those of nonlinear crystals when normalized to the active volumes [39, 40].

Although FE optical nonlinearities have mostly been observed in metals, analogous effects may also occur in heavily doped semiconductors. The dynamic properties of doped semiconductors undergo in fact an interesting transition from the size-quantization regime to the classical regime of plasmon oscillations. In the former regime, semiconductor quantum-wells have enabled the strongest nonlinear susceptibilities [41]. The resonant nature of electron transitions in quantum-wells however severely hinders their application in ultra-fast devices, since it inevitably causes a slower response and increased decoherence. The plasmonic regime for semiconductors, on the other hand, unlocks a new realm of possible ultra-fast nonlinear effects [42, 43], and offers several advantages with respect to nonlinear plasmonics in noble metals.

* cristian.ciraci@iit.it
To understand the advantage of using heavily doped semiconductors, it should be noticed that within the hydrodynamic description, third order nonlinear effects are inversely proportional to the square of the equilibrium free-carriers density, i.e. \( P_{NL}^{(3)} \propto n_0^{-1} \) (as we will show later), where \( P_{NL}^{(3)} \) is the third-order polarization vector. In gold, for example, the high concentration of free-carriers \( (n_0 \sim 10^{22} \text{ cm}^{-3}) \), leads to very weak contributions to the third-order nonlinear polarization if compared to contributions due to the crystal lattice nonlinear susceptibility \( \chi^{(3)} \). On the other hand, doped semiconductors with a plasma frequency in the MIR have a much lower charge density \( (n_0 \sim 10^{19} \text{ cm}^{-3}) \) than noble metals, such that nonlinear terms may grow as much as 6 order of magnitude, overcoming by far the lattice nonlinearities. Moreover, semiconductors have an effective electron mass, \( m \), that can be one order of magnitude smaller than that of noble metals. These characteristics cause also an increase of the nonlinear active volumes, whose measure (normalized to the operation wavelength) can be linked \([14, 15]\) to the Fermi velocity \( v_F \propto n_0^{1/3}/m \).

The large variety of semiconductors with different effective electron mass and doping levels give access to a wide range of possibilities to optimize and increase intrinsic nonlinear effects. Furthermore, these quantities, unlike in metals, can also be used to control their optical linear properties. Indeed, assuming a Drude model, a doped semiconductor is characterized by a dielectric function \( \varepsilon(\omega) = \varepsilon_{\infty} - \frac{\omega_p^2}{\omega^2 + i\gamma \omega} \), where \( \varepsilon_{\infty} \) is the background permittivity, \( \gamma \) is the damping rate and \( \omega_p = \sqrt{\frac{e^2 n_0}{\varepsilon_0 m}} \) is the plasma frequency of the material, being \( e \) and \( \varepsilon_0 \) the elementary charge (absolute value) and the dielectric constant of vacuum (note that hole-doping is unsuitable for plasmonics due to inter-valence-band transitions taking place in the MIR). By choosing the semiconductor, i.e. \( m \) and the doping level, \( n_0 \), it is then possible to tune the dielectric function at will \([16]\). For example, one could generate an epsilon-near-zero (ENZ) medium, to explore unconventional nonlinear optical effects \([26, 47]\). For all these reasons, the study of nonlinear FE dynamics in heavily doped semiconductors seems to represent a very promising test field toward the understanding of the fundamental mechanisms of nonlinearity in condensed matter and therefore in the direction of designing innovative optical technologies.

In this article, we focus on the process of third-harmonic generation (THG) from heavily doped semiconductors. We use a hydrodynamic description to study the free carriers dynamics up to third order terms, usually neglected for noble metals, and consider both cascaded and direct contributions. We first investigate THG efficiency in unpatterned doped semiconductor slabs. Interestingly, we find that cascaded THG due to SHG signals can be as strong as direct THG contributions even when the SHG efficiency is zero, i.e., there is no SHG in the far-field. We also suggest a possible experiment to demonstrate such effect. We then analyze the conditions that allow for the maximum THG efficiency and show that FE nonlinearities can be up to two orders of magnitude larger than intrinsic lattice nonlinearities. We finally introduce a grating pattern on the semiconductor slab surface that allows to strongly couple with normally incident waves and excite local plasmons to further enhance the conversion efficiency.

## II. HYDRODYNAMIC MODEL FOR THG

To model the many-body dynamics of interest, a theoretical framework that describes not only the charge density distribution and the nonlinear effects, but also nonlocal corrections, is needed. Indeed, because of the fermionic nature of the carriers, FE nonlinearities are intrinsically nonlocal, in the sense that the induced currents depend not only on the value of the electric field at a given point but also, through their spatial derivatives, on the value of the fields in the surrounding area. In addition, the formalism should give the possibility to compare FE nonlinearities to those arising from the crystal lattice, in order to solve a possible ambiguity in the origin of the nonlinear response. An appropriate representation of nonlinear and nonlocal FE dynamics in heavily doped semiconductors may be done taking into consideration the quasi-classical formalism of the hydrodynamic model, in the limit of Thomas-Fermi approximation \([14, 18, 19]\). The hydrodynamic theory has been successfully used to describe the FE dynamics in noble metal systems and is then extremely relevant to describe the behavior of heavily doped semiconductors. Furthermore, lattice nonlinearities can be easily included into the model by taking into consideration an intrinsic nonlinear susceptibility, as shown in the following.

Within the hydrodynamic formalism \([50, 52]\), the many-body dynamics of a FE fluid can be described taking into account two macroscopic variables, the charge density, \( n(r, t) \), and the current density, \( J(r, t) = -env \), where \( v(r, t) \) is the electron velocity field. Under the influence of external electric and magnetic fields, \( E(r, t) \) and \( H(r, t) \), the system can be described by the following equation:

\[
m \left( \frac{\partial}{\partial t} - \frac{J}{en} \cdot \nabla + \gamma \right) \frac{J}{en} = +eE - \frac{J}{n} \times \mu_0 H + \nabla \frac{\delta T_{TF}[n]}{\delta n},
\]

where \( \mu_0 \) is the magnetic permeability of vacuum. The last term takes into account the nonlocal effects due to the quantum pressure. Here, \( T_{TF}[n] \) is the kinetic energy functional in the Thomas-Fermi approximation, whose derivative with respect to \( n \) is \( \frac{\delta T_{TF}[n]}{\delta n} = c_{TF} \frac{5}{3} n^{2/3} \), with \( c_{TF} = \frac{\hbar^2}{m} \frac{1}{10} (3\pi^2)^{2/3} \) \([53]\).
Following a perturbative approach, it is possible to write\( n(r,t) = n_0 + n_{\text{ind}}(r,t) \), where \( n_{\text{ind}} = \frac{e}{\varepsilon} \mathbf{V} \cdot \mathbf{P} \) is the induced charge density and \( n_{\text{ind}} \ll n_0 \). Here \( \mathbf{P}(r,t) \) is the polarization vector, which is related to the current density through its time derivative, i.e. \( \dot{\mathbf{P}} = \mathbf{J} \). Eq. \( \text{(1)} \)
then becomes:
\[
\dot{\mathbf{P}} + \gamma \mathbf{P} = \frac{n_0 e^2}{m} \mathbf{E} + \beta^2 \nabla (\nabla \cdot \mathbf{P}) + S_{\text{NL}}^{(2)} + S_{\text{NL}}^{(3)}, \quad (2)
\]
where time derivatives are now expressed in dot notation. This equation is essentially constituted by the FE gas linear model, i.e. the Drude model plus a nonlocal linear term, where \( \beta^2 = \frac{1.3}{\varepsilon} \), and the second- and third-order nonlinear sources, \( S_{\text{NL}}^{(2)} \) and \( S_{\text{NL}}^{(3)} \), respectively, whose expressions are:
\[
S_{\text{NL}}^{(2)} = \frac{e}{m} \nabla \mathbf{\cdot} \mathbf{P} - \frac{e\mu_0}{m} \mathbf{\dot{P}} \times \mathbf{H} + \frac{1}{\varepsilon n_0} (\dot{\mathbf{P}} \nabla \cdot \mathbf{P} + \mathbf{P} \cdot \nabla \dot{\mathbf{P}})
+ \frac{1}{3} \frac{\beta^2}{\varepsilon n_0} (\nabla \cdot \mathbf{P})^2, \quad (3a)
\]
\[
S_{\text{NL}}^{(3)} = -\frac{1}{e^2 n_0^3} \left( \nabla \cdot \mathbf{P} \left( \dot{\mathbf{P}} \nabla \cdot \mathbf{P} + \mathbf{P} \cdot \nabla \dot{\mathbf{P}} \right) + \mathbf{P} \cdot \nabla \nabla \cdot \mathbf{P} \right)
- \frac{1}{2} \frac{\beta^2}{e n_0^2} (\nabla \nabla \cdot \mathbf{P})^3. \quad (3b)
\]
In deriving these equations, contributions up to the third order have been considered. In particular, we used \( n^{-1} \approx n_0^{-1} \left( 1 - \frac{n_{\text{ind}}}{n_0} \right) \) and \( n^{2/3} \approx n_0^{2/3} \left[ 1 + \frac{2}{3} \frac{n_{\text{ind}}}{n_0} - \frac{1}{3} \left( \frac{n_{\text{ind}}}{n_0} \right)^2 \right] \).
As it can be noted, the nonlinear sources include both local and nonlocal terms. In particular, second-order terms include Coulomb, Lorentz, convective and quantum pressure contributions in that order in Eq. \( \text{(3a)} \) while third-order terms in Eq. \( \text{(3b)} \) only include convective and pressure contributions.

In the most general picture, for a single input field, THG is the process in which a signal of frequency \( 3\omega \) is generated by the interaction of a field of frequency \( \omega \) with a nonlinear material. There are many ways, ideally infinite, to generate an output at \( 3\omega \). In the most common process, three photons of energy \( h \omega \) combine to give a single photon of energy \( 3h \omega \). In this case, the THG is usually referred to as direct THG. Additionally, the THG process can be seen as a two-steps process: first, a photon of energy \( 2h \omega \) is generated by the interaction of two photons of energy \( h \omega \); then, this combines to another photons of energy \( h \omega \) in order to get a photon of energy \( 3h \omega = 2h \omega + h \omega \). The THG is then the combination of two second-order nonlinear processes, namely SHG followed by sum-frequency generation and the process is commonly defined as cascaded THG. Being a combination of second-order processes, cascaded generation is forbidden in centrosymmetric bulk crystals, nevertheless, as we will show later, it is allowed in heavily doped semiconductors, similarly to metals \[\text{[54],[56]}\], where second-order nonlinearities arise from FE dynamics.

In order to describe the above mentioned process, let us assume time-harmonic dependence of the fields, i.e. \( F(r,t) = \sum F_j(r) e^{-i\omega_j t} \), with \( F = \mathbf{E}, \mathbf{H}, \text{or} \ \mathbf{P} \). Eq. \( \text{(2)} \) with Eqs. \( \text{(3)} \) and Maxwell’s equations can be rewritten as a set of equations for each harmonic \( \omega_j \) as follows:
\[
\nabla \nabla \cdot \mathbf{E}_j - \epsilon \omega_j^2 \frac{\mathbf{P}_j}{c^2} + \omega_j^2 \mu_0 (\mathbf{P}_j + \mathbf{P}_{\text{NL}}^{(\omega_j)}) = 0, \quad (4a)
\]
\[
\beta^2 (\nabla \cdot \mathbf{P}_j) + (\omega_j^2 + i\gamma \omega_j) \mathbf{P}_j = -\frac{n_0 e^2}{m} \mathbf{E}_j + \mathbf{S}_{\omega_j}. \quad (4b)
\]
In writing these equations we have considered dielectric local contributions from the semiconductor, both linear, through the local permittivity \( \varepsilon \), and nonlinear, through the nonlinear polarization \( \mathbf{P}_{\text{NL}}^{(\omega_j)} \). Coupling between different harmonics occurs through the nonlinear contributions \( \mathbf{P}_{\omega_j}^{\text{NL}} \) and \( \mathbf{S}_{\omega_j} \). For simplicity, we assume that the pump field is not affected by the nonlinear process (undepleted pump approximation), i.e. \( \mathbf{P}_{\omega_i}^{\text{NL}} = 0 \), as we expect harmonic signals to be several orders of magnitude smaller than the pump fields. The system of Eqs. \( \text{(4)} \) reduces then to three sets of one-way coupled equations, one for the fundamental \( (j = 1) \), one for the second-harmonic frequency \( (j = 2) \) and one for the third-harmonic frequency \( (j = 3) \).

The term that takes into account crystal lattice nonlinearities can be described using a bulk third-order susceptibility \( \chi^{(3)} \) as:
\[
\mathbf{P}_{\omega_3}^{\text{NL}} = \varepsilon_0 \chi^{(3)} (\mathbf{E}_1 \cdot \mathbf{E}_1) \mathbf{E}_1. \quad (5)
\]
In writing this equation we have assumed a centrosymmetric material, which also implies \( \chi^{(2)} = 0 \). The non-linear source terms due to free charges are:
\[
\mathbf{S}_{\omega_2} = -\frac{e}{m} (\mathbf{E}_1 \nabla \cdot \mathbf{P}_1) - i \frac{e\mu_0}{m} \omega_1 \mathbf{P}_1 \times \mathbf{H}_1 + \frac{\omega_1^2}{3 \varepsilon n_0} (\mathbf{P}_1 \cdot \nabla \cdot \mathbf{P}_1 + \mathbf{P}_1 \cdot \nabla \mathbf{P}_1) + \frac{2}{3} \frac{\beta^2}{\varepsilon n_0} (\nabla \cdot \mathbf{P}_1 \nabla \cdot \mathbf{P}_1), \quad (6)
\]
for the SHG and \( \mathbf{S}_{\omega_3} = \mathbf{S}_{\omega_2}^{(2)} + \mathbf{S}_{\omega_2}^{(3)} \) for the THG, with:
\[
\mathbf{S}_{\omega_2}^{(2)} = -\frac{e}{m} (\mathbf{E}_2 \nabla \cdot \mathbf{P}_1 + \mathbf{E}_1 \nabla \cdot \mathbf{P}_2) + \omega_2 \frac{e\mu_0}{m} (\mathbf{P}_2 \times \mathbf{H}_1 + \omega_1 \mathbf{P}_1 \times \mathbf{H}_2) + \frac{\omega_1^2 \omega_2}{3 \varepsilon n_0} (\mathbf{P}_2 \nabla \cdot \mathbf{P}_1 + \mathbf{P}_1 \nabla \cdot \mathbf{P}_2) + \mathbf{P}_1 \nabla \cdot \mathbf{P}_2 + \mathbf{P}_2 \cdot \nabla \mathbf{P}_2), \quad (7a)
\]
\[
\mathbf{S}_{\omega_2}^{(3)} = -\frac{\omega_1^2}{e^2 n_0^3} (\nabla \cdot \mathbf{P}_1 (\mathbf{P}_1 \nabla \cdot \mathbf{P}_1 + \mathbf{P}_1 \cdot \nabla \mathbf{P}_1) + \mathbf{P}_1 \cdot \nabla \nabla \cdot \mathbf{P}_1) + \frac{1}{2} \frac{\beta^2}{e^2 n_0^3} \nabla (\nabla \cdot \mathbf{P}_1)^3 \quad (7b)
\]
describing cascaded and direct THG due to FE dynamics, respectively. The hydrodynamic nonlinear sources in the general case of \( \mathbf{S}_{\omega_2} \neq 0 \) are reported in the appendix.
An initial step toward the understanding of nonlinear conversion in heavily doped semiconductors is to consider a very simple geometry, namely a slab of a semiconductor. Since a forcing electric field normal to the semiconductor surface is needed to induce a charge oscillation along the finite thickness of the slab, let us consider a transverse magnetic (TM) plane wave impinging on the slab with an angle of incidence, $\theta$, to induce a charge oscillation along the finite thickness of the slab.

By solving Eqs. (4-7) we are able to calculate second- and third-harmonic generated signals. We solved these equations numerically using the finite-element method (FEM) within a customized frequency-dependent implementation in COMSOL Multiphysics [57]. For our analysis, we consider Indium Phosphide (InP), a direct bandgap III-V semiconductor that has been recently studied as a low loss plasmonic material for the MIR region [58] [59], after being introduced among many alternative plasmonic materials few years ago [8]. InP is one of the most common materials in optoelectronic applications thanks to its easy integration and compatibility with the conventional microelectronic foundry and is particularly interesting for our study because of its very small effective electron mass, $m = 0.078 m_e$ [8], $m_e$ being the electron mass. MIR properties of undoped InP can be described through a constant linear permittivity $\varepsilon_{\infty} = 9.55$ [8] and a nonlinear bulk permittivity $\chi^{(3)} = 1.0 \times 10^{-18}$ m$^2$/V$^2$, a value that corresponds to the highest among most common semiconductors [60]. The pump peak intensity considered is $I_0 = 10$ W/µm$^2$ (1 GW/cm$^2$) in all cases and the THG efficiency has been obtained by normalizing the power of the generated signal to the input power at the fundamental frequency, $\eta = I_G/I_0$, where $I_G$ is the generated intensity. The simulated slab is 1 µm thick.

Figure 1 shows SHG and THG efficiencies as a function of the incident angle, $\theta$, for a InP slab with doping level of $n_0 = 5.0 \times 10^{19}$ cm$^{-3}$ excited by a fundamental field (FF) oscillating at $\lambda_{FF} = 10$ µm. As one would expect for a centrosymmetric material, the SHG signal at normal incidence is zero while it peaks at larger angles, $\theta \simeq 72^\circ$, similarly to SHG from noble metals [61][62]. The THG efficiency, on the other hand, gives a non-zero contribution also at normal incidence, while increasing for large angles, peaking at $\theta \simeq 77^\circ$. As it can be inferred by Eq. (7b), direct FE contribution are zero at normal incidence, since $\nabla \cdot \mathbf{P}_1$ is identically zero for $\theta = 0$. Hence, it would be logical to think that the contribution shown in Fig. 1d are due to the lattice nonlinearities, since also the SHG is zero. In order to clarify this point, in Fig. 1 we plot the THG efficiencies associated to each contributions. Surprisingly the FE cascaded THG signal is not zero at normal incidence, moreover it is comparable to the lattice $\chi^{(3)}$ contribution. How is it possible to obtain a cascaded effect from a signal that is apparently zero? To understand this, we plot in Fig. 1 the non-zero component of the fundamental and second-harmonic field in the vicinity of the semiconductor surface. It can be observed indeed that the second-harmonic field has a non-zero longitudinal component that follows the profile of the square FF amplitude. This component is induced by the Lorentz term in Eq. (6) inside the semiconductor. However, because the vacuum does not support longitudinal modes, such component cannot be coupled to the far-field. On the other hand, a second-harmonic longitudinal component in the semiconductor combines back with the magnetic field associated to the FF (through the Lorentz term in Eq. (7a)), finally resulting in a transverse electric field component at the third harmonic. A similar contribution is also given by the one of the convective terms. Indeed, a THG cascaded contribution can be generated through an apparently zero SHG.

Although this effect can be easily observed with numerical simulations, it is not that simple to differentiate between cascaded and direct contributions in an experimental set up. Let us anticipate however that it is indeed possible to take advantage from this process and experimentally demonstrate this peculiar phenomenon by properly tuning the frequency with respect to the natural ENZ frequency of doped semiconductors, as we will show below.
In Fig. 2, each map plot reports the simulated reflected THG efficiency $\eta$, of a slab of doped semiconductor as function of the FF incident angle $\theta$, and wavelength $\lambda_{FF}$, for different conditions. Specifically, each row refers to the THG efficiency obtained considering a distinct nonlinear source: $S_{nl}^{(3)}$ (FE direct), $S_{nl}^{(2)}$ (FE cascaded) and $\omega^2\mu_0\mathcal{P}_{NL}$ (lattice) respectively. Each column refers to a different level of doping, and, as a consequence, to a different screened plasma frequency, $\tilde{\omega}_p \simeq \omega_p/\sqrt{\varepsilon_{\infty}}$.

As discussed in the introduction, a higher $n_0$ is predicted to correspond to smaller hydrodynamic effects and thus to a smaller THG efficiency for a given ratio $\omega_{FF}/\omega_p$. For instance, if one considers the FE direct contribution, when $\lambda_{FF} = \tilde{\lambda}_p = \frac{2\omega_p}{\omega_{FF}}$, $\eta$ goes from a peak of the order of $10^{-8}$ in Fig. 2a, to a maximum of $10^{-6}$ in Fig. 2f, i.e. the efficiency is between 10 and 100 times bigger for a doping five times smaller. Instead, for a given level of doping, an increase of the efficiency is expected when there are the most favourable energetic condition for the enhancement of the fields involved in the process. Peaks of the efficiency are hence predicted for $\lambda_{FF} = \tilde{\lambda}_p$ and $\lambda_{FF} = 3\tilde{\lambda}_p$ in all cases, namely when the FF and the generated field are at ENZ condition, respectively. These features can be clearly seen in all plots of Fig. 2 when the FF is at the ENZ condition. However, when the generated signal is at the ENZ condition (only possible for the first column, for $\lambda_{FF} \simeq 12 \mu m$), there is a sharp peak only for the FE direct THG (Fig. 2a). In the case of the cascaded process, there is no a clear maximum of $\eta$, probably because an enhancement of the generated field is not so effective for the overall efficiency if the other fields involved in the

B. Hydrodynamic vs crystalline lattice contributions

Having clarified the nature of cascaded effects, let us now focus on spectral properties of the THG process. For a given material characterized by an effective mass $m$ and a fixed doping level $n_0$, the nature of its optical response depends on the frequency of the external electromagnetic field. That is, the doped semiconductor has a metallic behavior for frequencies below the plasma frequency, while it behaves like a dielectric for higher frequencies.

In Fig. 2, each map plot reports the simulated reflected THG efficiency, $\eta$, of a slab of doped semiconductor as function of the FF incident angle $\theta$, and wavelength, $\lambda_{FF}$, for different conditions. Specifically, each row refers to the THG efficiency obtained considering a distinct nonlinear source: $S_{nl}^{(3)}$ (FE direct), $S_{nl}^{(2)}$ (FE cascaded) and $\omega^2\mu_0\mathcal{P}_{NL}$ (lattice) respectively. Each column refers to a different level of doping, and, as a consequence, to a different screened plasma frequency, $\tilde{\omega}_p \simeq \omega_p/\sqrt{\varepsilon_{\infty}}$.

As discussed in the introduction, a higher $n_0$ is predicted to correspond to smaller hydrodynamic effects and thus to a smaller THG efficiency for a given ratio $\omega_{FF}/\omega_p$. For instance, if one considers the FE direct contribution, when $\lambda_{FF} = \tilde{\lambda}_p = \frac{2\omega_p}{\omega_{FF}}$, $\eta$ goes from a peak of the order of $10^{-8}$ in Fig. 2a, to a maximum of $10^{-6}$ in Fig. 2f, i.e. the efficiency is between 10 and 100 times bigger for a doping five times smaller. Instead, for a given level of doping, an increase of the efficiency is expected when there are the most favourable energetic condition for the enhancement of the fields involved in the process. Peaks of the efficiency are hence predicted for $\lambda_{FF} = \tilde{\lambda}_p$ and $\lambda_{FF} = 3\tilde{\lambda}_p$ in all cases, namely when the FF and the generated field are at ENZ condition, respectively. These features can be clearly seen in all plots of Fig. 2 when the FF is at the ENZ condition. However, when the generated signal is at the ENZ condition (only possible for the first column, for $\lambda_{FF} \simeq 12 \mu m$), there is a sharp peak only for the FE direct THG (Fig. 2a). In the case of the cascaded process, there is no a clear maximum of $\eta$, probably because an enhancement of the generated field is not so effective for the overall efficiency if the other fields involved in the

B. Hydrodynamic vs crystalline lattice contributions

Having clarified the nature of cascaded effects, let us now focus on spectral properties of the THG process. For a given material characterized by an effective mass $m$ and a fixed doping level $n_0$, the nature of its optical response depends on the frequency of the external electromagnetic field. That is, the doped semiconductor has a metallic behavior for frequencies below the plasma frequency, while it behaves like a dielectric for higher frequencies.

In Fig. 2, each map plot reports the simulated reflected THG efficiency, $\eta$, of a slab of doped semiconductor as function of the FF incident angle $\theta$, and wavelength, $\lambda_{FF}$, for different conditions. Specifically, each row refers to the THG efficiency obtained considering a distinct nonlinear source: $S_{nl}^{(3)}$ (FE direct), $S_{nl}^{(2)}$ (FE cascaded) and $\omega^2\mu_0\mathcal{P}_{NL}$ (lattice) respectively. Each column refers to a different level of doping, and, as a consequence, to a different screened plasma frequency, $\tilde{\omega}_p \simeq \omega_p/\sqrt{\varepsilon_{\infty}}$.

As discussed in the introduction, a higher $n_0$ is predicted to correspond to smaller hydrodynamic effects and thus to a smaller THG efficiency for a given ratio $\omega_{FF}/\omega_p$. For instance, if one considers the FE direct contribution, when $\lambda_{FF} = \tilde{\lambda}_p = \frac{2\omega_p}{\omega_{FF}}$, $\eta$ goes from a peak of the order of $10^{-8}$ in Fig. 2a, to a maximum of $10^{-6}$ in Fig. 2f, i.e. the efficiency is between 10 and 100 times bigger for a doping five times smaller. Instead, for a given level of doping, an increase of the efficiency is expected when there are the most favourable energetic condition for the enhancement of the fields involved in the process. Peaks of the efficiency are hence predicted for $\lambda_{FF} = \tilde{\lambda}_p$ and $\lambda_{FF} = 3\tilde{\lambda}_p$ in all cases, namely when the FF and the generated field are at ENZ condition, respectively. These features can be clearly seen in all plots of Fig. 2 when the FF is at the ENZ condition. However, when the generated signal is at the ENZ condition (only possible for the first column, for $\lambda_{FF} \simeq 12 \mu m$), there is a sharp peak only for the FE direct THG (Fig. 2a). In the case of the cascaded process, there is no a clear maximum of $\eta$, probably because an enhancement of the generated field is not so effective for the overall efficiency if the other fields involved in the
FIG. 3. (a) THG efficiency $\eta$ of a highly doped InP slab as a function of the incident angle $\theta$ and of the wavelength $\lambda_{FF}$ of the FF, calculated for $\lambda_p \approx 4$ ($n_0 = 5.0 \times 10^{19}$ cm$^{-3}$) taking into account all the nonlinear sources; (b) THG efficiency as a function of the background permittivity for the three distinct nonlinear contributions, with $n_0 = 5.0 \times 10^{19}$ cm$^{-3}$, $m = 0.1 \, m_e$, $\lambda_{FF} = 1.8 \, \lambda_p$ and $\theta = 40^\circ$.

process are not enhanced as well. The same happens for the lattice nonlinearities because of the limited interaction of the input field with the slab in the metallic regime of the material, since these contributions originate in the bulk.

A maximum of the efficiency for $\lambda_{FF} = 2\lambda_p$ is by contrast a unique feature expected only for the cascaded THG. It corresponds to the enhancement of the field generated through SHG when it is at the ENZ condition. The aforementioned peculiarity represents a crucial clue for the experimental demonstration of cascade effects in heavily doped semiconductors and can be clearly seen in Figs. 2d and 2e, for $\lambda_p \approx 8$ $\mu$m and $\lambda_{FF} \approx 12$ $\mu$m, respectively.

A very important characteristic that emerges from Fig. 2 is the order of magnitude of the FE THG compared to that arising from lattice nonlinearities. Indeed, although we used a very high value of $\chi^{(3)}$, third-order hydrodynamic contributions can be the most important in the portion of the MIR we considered. This could be expected in the metallic regime of the material far from the screened plasma wavelength, where the interaction of the pump with the bulk crystal lattice is hindered. However, for the smaller level of doping the efficiency of THG is higher for the hydrodynamic direct process than for the lattice even at ENZ. This represents a remarkable prediction toward the application of heavily doped semiconductor for nonlinear plasmonics.

C. What to expect from experiments

At this point of the discussion, it is interesting to wonder what kind of results one should expect from an experiment. Let us now consider all the nonlinear sources at once to show what it is expected in a more realistic setup. The results are reported in Fig. 3, in the case of $\lambda_p \approx 4$ $\mu$m ($n_0 = 5.0 \times 10^{19}$ cm$^{-3}$). Unsurprisingly, in Fig. 3 one can see all the main characteristics of the three nonlinear sources, which appear separately in the first column of Fig. 2 combined in a single plot. If our model is accurate enough in describing nonlinearities in heavily doped semiconductors, an interesting way to demonstrate our predictions, taking into consideration Fig. 3a, seems to look for a very peculiar trend of the THG efficiency at wavelengths close to $2\lambda_p$ and to $3\lambda_p$. Comparing Fig. 3a and Fig. 2b, it is clear that the peak at $2\lambda_p$ can only be caused by a cascaded generation. A similar reasoning seems to make sense for the peak at $3\lambda_p$ when considering the FE direct THG.

D. The influence of the semiconductor background permittivity

Hitherto, we considered a well-defined material, however, it is interesting to study how the hydrodynamic nonlinear sources are influenced by the material physical properties other the doping level $n_0$, even if the actual physical parameter cannot be varied in the experiments, as it is the case for the background permittivity, $\varepsilon_\infty$. In Fig. 3b we show the THG efficiency, $\eta$, as function of the $\varepsilon_\infty$ for the three THG contributions, with all other variables unchanged. The trend of hydrodynamic THG efficiencies in Fig. 3b might be intuitively understood thinking to the interaction of light with the slab from the point of view of Snell’s law. Indeed, the refractive index increases with the square root of $\varepsilon_\infty$. This also defines the angle of refraction of a light inside a material (the higher is the refractive index, the smaller is the angle of refraction). Therefore, for a given angle of incidence, a higher $\varepsilon_\infty$ causes the transmitted FF at the interface air-semiconductor to be closer to normal incidence when propagating inside the slab. However, we know from the previous paragraphs that surface contributions are hindered at normal incidence, so, the FE direct THG efficiency, due only to surface contributions (see Eq. (7a)) becomes smaller when increasing $\varepsilon_\infty$. Conversely, the Lorentz term is maximum at normal incidence and it is then favoured for higher $\varepsilon_\infty$. As a consequence, taking into account Eq. (7a), the cascaded THG efficiency has a very peculiar trend caused by a balance between surface and Lorentz contributions. Indeed, in Fig. 3b, the FE cascaded contribution curve has the same decreasing trend as the FE direct contribution one for very small values of $\varepsilon_\infty$, i.e. far from normal refractive, where surface contributions are dominant with respect to the Lorentz term. It reaches a minimum where the contributions considered are of the same order and finally grows steadily with $\varepsilon_\infty$ when the Lorentz contribution becomes the most relevant (closer to normal refraction). Instead, for background lattice nonlinearities, $\eta$ is higher for lower $\varepsilon_\infty$, simply because of the higher transparency of the material. As a consequence, one should expect relatively higher hydrodynamic nonlinear contributions when considering a material with a lower background permittivity such as the conductive Indium Tin-Oxide (ITO) [49].
The analysis developed so far was useful to understand some fundamental properties of hydrodynamic nonlinearities. From the point of view of technology, however, it is perhaps more interesting to apply our model to a nanopatterned semiconductor slab with a geometrical resonance in the MIR. To this aim, we take now into consideration an infinite array of parallel subwavelength grooves (a grating), as depicted in Fig. 4a. This system supports plasmonic resonances and allows to couple virtually all incident energy into the active material \( \text{InP} \), \( \text{InAs} \), \( \text{GaAs} \), \( \text{GaP} \), \( \text{GaSb} \), \( \text{MgO} \), \( \text{SiO}_2 \), \( \text{B}_{2}\text{O}_3 \). The peak efficiency of FE contributions is several order of magnitude higher than in the case of the slab and it always surpasses that of intrinsic lattice nonlinearities. The maximum value was around \( 10^{-6} \) for the direct process and smaller for the cascaded, instead, here we get a value larger than \( 10^{-4} \) for the first process and bigger than \( 10^{-3} \) for the cascaded. Consequently, for the grooves array the efficiency is at least \( 10^2 \) times larger for the FE direct THG and more than \( 10^4 \) times higher for the cascaded THG, with the same pump intensity; while the THG due to the lattice increases just by a factor of 10. Therefore, although we have considered a material with very high third-order nonlinear susceptibility, \( \chi^{(3)} \), we predict that FE contributions will overcome by far conventional semiconductor nonlinearities in THG processes in the MIR.

V. CONCLUSIONS AND PERSPECTIVES

We have used a hydrodynamic model to describe FE nonlinear optical dynamics in heavily doped semiconductors. In particular, by deriving nonlinear hydrodynamic terms up to the third-order, we have investigated FE contributions to the THG process. Contrarily to noble metals at optical frequencies, heavily doped semiconductor third-harmonic response is predicted to be strongly driven by hydrodynamic nonlinearities. We have found in fact that such nonlinearities might generated THG signals up to two orders of magnitude larger than conventional semiconductor nonlinearities when combined to plasmonic field enhancement. Moreover, we have showed that cascaded contributions, which are very often neglected, might be extremely important, even when one of

\[ \chi^{(3)} \]

\[ \text{InP} \]

\[ \text{InAs} \]

\[ \text{GaAs} \]

\[ \text{GaP} \]

\[ \text{GaSb} \]

\[ \text{MgO} \]

\[ \text{SiO}_2 \]

\[ \text{B}_{2}\text{O}_3 \]
the constituent signals appear to be negligible. Although in most realistic experiments it would be impossible to
differentiate cascaded effects from conventional nonlinear
signals, we have proposed a simple method to experimen-
tally verify our findings. Field-effect gated devices and
optical generation of free carriers, both characteristics of
semiconductors, could provide a route for tunable non-
linear optical systems on semiconductor-based photonic
integrated circuit platforms. We believe that these re-
results could open new avenues for integrated nonlinear
optics at MIR frequencies and beyond.

Appendix A: Hydrodynamic nonlinear sources

In section [I] we derived the system of equations Eq. [3] and, assuming $S_{\omega_1} = 0$, we reported in Eqs. [6] only the
nonlinear sources $S_{\omega_2}$ and $S_{\omega_3}$. However, in the most general case $S_{\omega_1} \neq 0$. It means that, for high input intensities, the nonlinear field produced may be converted back to the FF. Taking into account Eqs. [3] this may happen through difference-frequency generation of the second order:

$$S^{(2)}_{\omega_1=\omega_2-\omega_1} = -\frac{e}{m} \left( E_1 \nabla \cdot P_2 + E_2 \nabla \cdot P_1^* \right) - \frac{\epsilon \mu_0}{m} (\omega_2 P_2 \times H_1 - \omega_1 P_1^* \times H_2) + \frac{\omega_1^2}{en_0} (P_2 \nabla \cdot P_1^* + P_2 \nabla P_1^* + P_1^* \nabla \cdot P_2 + P_1^* \nabla P_2) - \frac{2\beta^2}{3en_0} \left( \nabla \cdot P_2 \nabla \cdot P_1^* + \nabla \cdot P_1^* \nabla \cdot P_2 \right),$$

(A1)

through difference-frequency generation of the third order:

$$S^{(3)}_{\omega_1=\omega_3-\omega_1-\omega_1} = -\frac{\omega_1^2}{en_0^2} \left[ (\nabla \cdot P_3) (P_1^* \nabla \cdot P_1^* + P_1^* \nabla P_1^*) + P_1^* \nabla \cdot P_3 \nabla (\nabla \cdot P_3) \right] + \frac{\omega_1^2}{en_0^2} \left[ (\nabla \cdot P_1^*) (P_3 \nabla \cdot P_1^* + P_3 \nabla P_1^* + P_1^* \nabla \cdot P_3 + P_1^* \nabla P_3) + 2P_3 \cdot P_1^* \nabla (\nabla \cdot P_1^*) \right] + \frac{1}{9} \frac{\beta^2}{en_0^2} \nabla \left[ (\nabla \cdot P_1^*)^2 (\nabla \cdot P_3) \right],$$

(A2)

eventually, third order effects may also occur as Kerr type nonlinearities:

$$S^{(3)}_{\omega_1=\omega_1-\omega_1+\omega_1} = -\frac{\omega_1^2}{en_0^2} \left[ (\nabla \cdot P_1^*) (P_1 \cdot \nabla P_1) + P_1 \cdot P_1 \nabla (\nabla \cdot P_1^*) \right] + \frac{\omega_1^2}{en_0^2} \left[ (\nabla \cdot P_1) (P_1 \cdot \nabla P_1^* + P_1^* \nabla \cdot P_1 + P_1^* \nabla P_1) + 2 | P_1 |^2 \nabla (\nabla \cdot P_1) \right] + \frac{1}{9} \frac{\beta^2}{en_0^2} \nabla \left[ (\nabla \cdot P_1) |^2 (\nabla \cdot P_1) \right],$$

(A3)

where all the variables and the constants have been introduced in sections [I] and [II]. As described in section [I] in deriving
these equations we assumed time-harmonic dependence of the fields.

[1] J. J. Baumberg, J. Aizpurua, M. H. Mikkelsen, and D. R.
Smith, Extreme nanophotonics from ultrathin metallic
gaps, Nat. Mater. 18, 668 (2019).
[2] O. S. Ojambati, R. Chikkaraddy, W. D. Deacon, M. Hor-
ton, D. Kos, V. A. Turek, U. F. Keyser, and J. J. Baumberg,
Quantum electrodynamics at room temperature
coupling a single vibrating molecule with a plasmonic
nanocavity, Nat. Commun. 10, 1049 (2019).
[3] J. Huang, A. J. Traverso, G. Yang, and M. H. Mikkelsen,
Real-Time Tunable Strong Coupling: From Individual
Nanocavities to Metasurfaces, ACS Photonics 6, 838
(2019).
[4] W. Chen, S. Zhang, M. Kang, W. Liu, Z. Ou, Y. Li,
Y. Zhang, Z. Guan, and H. Xu, Probing the limits of
plasmonic enhancement using a two-dimensional atomic
crystal probe, Light: Science & Applications 7, 1 (2018).
[5] A. I. Fernández-Domínguez, S. I. Bozhevolnyi, and N. A.
Mortensen, Plasmon-Enhanced Generation of Nonclassical
Light, ACS Photonics 5, 3447 (2018).
[6] A. F. Koenderink, Single-Photon Nanoantennas, ACS
Photonics 4, 710 (2017).
[7] A. Boltasseva and H. A. Atwater, Low-Loss Plasmonic
Metamaterials, Science 331, 290 (2011).
[8] G. V. Naik, V. M. Shalaev, and A. Boltasseva, Alterna-
tive plasmonic materials: Beyond gold and silver, Ad-
vanced Materials 25, 3264 (2013).
[10] Y. Su, W. Wang, X. Hu, H. Hu, X. Huang, Y. Wang, J. Si, X. Xie, B. Han, H. Feng, Q. Hao, G. Zhu, T. Duan, and W. Zhao, 10 Gbps DPSK transmission over free-space link in the mid-infrared, Opt. Express 26, 34515 (2018).

[11] K. Koshelev, S. Kruk, E. Melik-Gaykazyan, J.-H. Choi, A. Bogdanov, H.-G. Park, and Y. Kivshar, Subwavelength dielectric resonators for nonlinear nanophotonics, Science 367, 288 (2020).

[12] V. F. Gili, L. Carletti, A. Locatelli, D. Rocco, M. Finazzi, L. Ghirardini, I. Favero, C. Gomez, A. Lemaître, M. Celibran, C. De Angelis, and G. Leo, Monolithic AlGaAs second-harmonic nanoantennas, Opt. Express 24, 15965 (2016).

[13] L. Carletti, G. Marino, L. Ghirardini, V. F. Gili, D. Rocco, I. Favero, A. Locatelli, A. V. Zayats, M.Celebrano, M. Finazzi, G. Leo, C. de Angelis, and D. N. Neshev, Nonlinear Goniometry by Second-Harmonic Generation in AlGaAs Nanonanotennas, ACS Photonics 5, 4386 (2018).

[14] W. J. G. Y. A. W. R. Qixin Shen and M. H. Mikkelsen, Active Control of Multiple, Simultaneous Nonlinear Optical Processes in Plasmonic Nanopag Cavities, ACS Photonics 7, 1 (2020).

[15] R. Sarma, D. de Ceglia, N. Nookala, M. A. Vincenti, S. Campione, O. Wolf, M. Scalora, M. B. Sinclair, M. A. Belkin, and I. Brener, Broadband and Efficient Second-Harmonic Generation from a Hybrid Dielectric Metasurface/ Semiconductor Quantum-Well Structure, ACS Photonics 6, 1458 (2019).

[16] Y. Zeng, H. Qian, M. J. Rozin, Z. Liu, and A. R. Tao, Enhanced Second Harmonic Generation in Double-Resonance Colloidal Metasurfaces, Adv. Funct. Mater. 28, 1803019 (2018).

[17] A. R. Echarri, J. D. Cox, R. Yu, and F. J. G. de Abajo, Enhancement of nonlinear optical phenomena by localized resonances, ACS Photonics 5, 1521 (2018).

[18] Q. Shen, T. B. Hoang, G. Yang, V. D. Wheeler, and M. H. Mikkelsen, Probing the origin of highly-efficient third-harmonic generation in plasmonic nanogaps, Opt. Express 26, 20718 (2018).

[19] S. Guddala and S. A. Ramakrishna, Optical limiting by nonlinear tuning of resonance in metamaterial absorbers, Opt. Lett. 41, 5150 (2016).

[20] T. Shibamuna, G. Grinblat, P. Albella, and S. A. Maier, Efficient Third Harmonic Generation from Metal – Dielectric Hybrid Nanoantennas, Nano Lett. 17, 6472651 (2017).

[21] J. Lee, M. Tymchenko, C. Argyropoulos, P.-Y. Chen, F. Lu, F. Demmerle, G. Boehm, M.-C. Amann, A. Ali, and M. A. Belkin, Giant nonlinear response from plasmonic metasurfaces coupled to intersubband transitions, Nature 511, 65 (2014).

[22] F. Wang, A. B. F. Martinson, and H. Harutyunyan, Efficient Nonlinear Metasurface Based on Nonplanar Plasmonic Nanocavities, ACS Photonics 4, 1188 (2017).

[23] A. Noor, A. R. Damodaran, I.-H. Lee, S. A. Maier, S.-H. Oh, and C. Ciraci, Mode-Matching Enhancement of Second-Harmonic Generation with Plasmonic Nanopatch Antennas, ACS Photonics 1, 1 (2020).

[24] M. Hentschel, B. Metzger, B. Knabe, K. Buse, and H. Giessen, Linear and nonlinear optical properties of hybrid metallic–dielectric plasmonic nanoantennas, Beilstein J. Nanotechnol. 7, 111 (2016).

[25] E. Barakat, M.-P. Bernal, and F. I. Baida, Theoretical analysis of enhanced nonlinear conversion from metallo-dielectric nano-structures, Opt. Express 20, 16258 (2012).

[26] J. Deng, Y. Tang, S. Chen, K. Li, A. V. Zayats, and G. Li, Giant Enhancement of Second-Order Nonlinearity of Epsilon-near- Zero Medium by a Plasmonic Metasurface, Nano Lett. 20, 5421 (2020).

[27] M. P. Nielsen, X. Shi, P. Dichtl, S. A. Maier, and R. F. Oulton, Giant nonlinear response at a plasmonic nano-focus drives efficient four-wave mixing, Science 358, 1179 (2017).

[28] H. Harutyunyan, G. Volpe, R. Quidant, and L. Novotny, Enhancing the Nonlinear Optical Response Using Multifrequency Gold-Nanowire Antennas, Phys. Rev. Lett. 108, 217403 (2012).

[29] K. D. Ko, A. Kumar, K. H. Fung, R. Ambekar, G. L. Liu, N. X. Fang, and K. C. Toussaint, Nonlinear optical response from arrays of Au bowtie nanoantennas, Nano Lett. 11, 61 (2011).

[30] R. Czaplicki, A. Kiviniemi, M. J. Huttunen, X. Zang, T. Stolt, I. Vartiainen, J. Butet, M. Kuitinen, O. J. F. Martin, and M. Kauranen, Less Is More: Enhancement of Second-Harmonic Generation from Metasurfaces by Reduced Nanoparticle Density, Nano Lett. 18, 7709 (2018).

[31] A. V. Krasavin, P. Ginzburg, and A. V. Zayats, Free-electron Optical Nonlinearities in Plasmonic Nanostructures: A Review of the Hydrodynamic Description, Laser & Photon. Rev. 12, 1700082 (2018).

[32] A. Chizmeshya and E. Zaremba, Second-harmonic generation at metal surfaces using an extended Thomas Fermi von Weizsacker theory, Phys. Rev. B 37, 2805 (1988).

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

[34] 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).

[35] Y. Pavlyukh, J. Berakdar, and W. Hubner, Semi-classical approximation for second-harmonic generation in nanoparticles, New J. Phys. 14, 093044 (2012).

[36] C. Ciraci, E. Poutrina, M. Scalora, and D. R. Smith, Origin of second-harmonic generation enhancement in optical split-ring resonators, Phys. Rev. B 85, 214034(R) (2012).

[37] A. Cappetti, C. Forestiere, L. Dal Negro, and G. Miano, Full-Wave Analytical Solution of Second-Harmonic Generation in Metal Nanospheres, Plasmonics 9, 151 (2013).

[38] B. Wells, A. Y. Bykov, G. Marino, M. E. Nasir, A. V. Zayats, and V. A. Podolskiy, Structural second-order nonlinearity in plasmonic metamaterials, Optica 5, 1502 (2018).

[39] M. W. Klein, M. Wegener, N. Feth, and S. Linden, Experiments on second- and third-harmonic generation from magnetic metamaterials, Opt. Express 15, 5238 (2007).

[40] R. Czaplicki, H. Husu, R. Siikanen, J. Mäkitalo, M. Kauranen, J. Laukkanen, J. Lehtolahti, and M. Kuittinen, Enhancement of Second-Harmonic Generation from Metal Nanoparticles by Passive Elements, Phys. Rev. Lett. 110, 093902 (2013).
[41] J. Lee, M. Tymchenko, C. Argyropoulos, P.-Y. Chen, F. Lu, F. Demmerle, G. Boehm, M.-C. Amann, A. Ali, and M. A. Belkin, Giant nonlinear response from plasmonic metasurfaces coupled to intersubband transitions, Nature 511, 65 (2014).

[42] M. P. Fischer, A. Riede, K. Gallacher, J. Frigerio, G. Pellegrini, M. Ortolani, D. J. Paul, G. Isella, A. Leitenstorfer, P. Biagioni, and D. Brida, Plasmonic mid-infrared third harmonic generation in germanium nanoantennas, Light: Science & Applications 7, 1 (2018).

[43] M. Wagner, A. S. McLeod, S. J. Maddox, Z. Fei, M. Liu, R. D. Averitt, M. M. Fogler, S. R. Bank, F. Keilmann, and D. N. Basov, Ultrafast dynamics of surface plasmons in InAs by time-resolved infrared nanospectroscopy, Nano Letters 14, 4529 (2014). https://doi.org/10.1021/nl501558t.

[44] J. R. Maack, N. A. Mortensen, and M. Wubs, Size-dependent nonlocal effects in plasmonic semiconductor nanoparticles, EPL 119, 17003 (2017).

[45] E. J. C. Dias, D. A. Iranzo, P. A. D. Gonçalves, Y. Hajati, Y. V. Bludov, A.-P. Jauho, N. A. Mortensen, F. H. L. Koppens, and N. M. R. Peres, Probing nonlocal effects in metals with graphene plasmons, Phys. Rev. B 97, 245405 (2018).

[46] J. Frigerio, A. Ballabio, G. Isella, E. Sakat, G. Pellegrini, P. Biagioni, M. Bollani, E. Napolitani, C. Manganelli, M. Virgilio, A. Grupp, M. P. Fischer, D. Brida, K. Gallacher, D. J. Paul, L. Baldassarre, P. Calvani, V. Gliberti, A. Nucara, and M. Ortolani, Tunability of the dielectric function of heavily doped germanium thin films for mid-infrared plasmonics, Phys. Rev. B 94, 085202 (2016).

[47] O. Reshef, I. De Leon, M. Z. Alam, and R. W. Boyd, Nonlinear optical effects in epsilon-near-zero media, Nat. Rev. Mater. 4, 535 (2019).

[48] M. Scalora, J. Trull, D. de Ceglia, M. A. Vincenti, N. Akozbek, Z. Copps, L. Rodríguez-Suné, and C. Cojocaru, Electrodynamic properties of conductive oxides: Intensity-dependent anisotropy, reconstruction of the effective dielectric constant, and harmonic generation, Phys. Rev. A 101, 053828 (2020).

[49] L. Rodríguez-Suné, M. Scalora, A. S. Johnson, C. Cojocaru, N. Akozbek, Z. J. Copps, D. Perez-Salinas, S. Wall, and J. Trull, Study of second and third harmonic generation from an indium tin oxide nanolayer: Influence of nonlocal effects and hot electrons, APL Photon. 5, 010801 (2020).

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

[51] D. de Ceglia, M. Scalora, M. A. Vincenti, S. Campione, K. Kelley, E. L. Runnerstrom, J.-P. Maria, G. A. Keeler, and T. S. Luk, Viscoelastic optical nonlocality of low-loss epsilon-near-zero nanofilms, Sci. Rep. 8, 874 (2018).

[52] S. Raza, S. I. Bozhevolnyi, M. Wubs, and N. A. Mortensen, Nonlocal optical response in metallic nanostructures, J. Phys.: Condens. Mat. 27, 183204 (2015).

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

[54] M. Celebrano, X. Wu, M. Baselli, S. G. mann, P. Baggioini, A. Locatelli, C. de Angelis, G. Cerullo, R. Osel-lame, B. Hecht, L. Duò, F. Cicciacci, and M. Finazzi, Mode matching in multiresonant plasmonic nanostructures for enhanced second harmonic generation, Nature Nanotechnology 10, 1 (2015).

[55] F. De Luca and C. Ciracì, Difference-frequency generation in plasmonic nanostructures: a parameter-free hydrodynamic description, J. Opt. Soc. Am. B 36, 1979 (2019).

[56] M. Celebrano, A. Locatelli, L. Ghirardini, G. Pellegrini, P. Biagioni, A. Zilli, X. Wu, S. Grossmann, L. Carletti, C. De Angelis, L. Duò, B. Hecht, and M. Finazzi, Evidence of cascaded third-harmonic generation in noncentrosymmetric gold nanoantennas, Nano Lett. 19, 7013 (2019).

[57] COMSOL Multiphysics, www.comsol.com.

[58] M. A. Panah, O. Takayama, S. V. Morozov, K. E. Kudryavtsev, E. S. Semenova, and A. V. Lavrinenko, Highly doped InP as a low loss plasmonic material for mid-ir region, Opt. Express 24, 29077 (2016).

[59] M. A. Panah, L. Han, K. Norrman, N. Pryds, A. Nadtochiy, A. Zhukov, A. V. Lavrinenko, and E. S. Semenova, Mid-ir optical properties of silicon doped InP, Opt. Mater. Express 7, 2260 (2017).

[60] R. W. Boyd, Nonlinear Optics (Academic Press, San Diego, 2006).

[61] K. A. O’Donnell and R. Torre, Characterization of the second-harmonic response of a silver–air interface, New J. Phys. 7, 154 (2005).

[62] L. Rodríguez-Suné, J. Trull, C. Cojocaru, N. Akozbek, D. de Ceglia, M. A. Vincenti, and M. Scalora, Harmonic generation from gold nanolayers: an old problem under a new light, arXiv:1904.09279, 1 (2020).

[63] J. Le Perchec, P. Quémerais, A. Barbara, and T. López-Ríos, Why Metallic Surfaces with Grooves a Few Nanometers Deep and Wide May Strongly Absorb Visible Light, Phys. Rev. Lett. 100, 066408 (2008).

[64] A. Pitelet, N. Schmitt, D. Loukrezis, C. Scheid, H. De Gersem, C. Ciraci, E. Centeno, and A. Moreau, Influence of spatial dispersion on surface plasmons, nanoparticles, and grating couplers, J. Opt. Soc. Am. B 36, 2989 (2019).

[65] M. Dechaux, P. H. Tichit, C. Ciraci, J. Benedicto, R. Polles, E. Centeno, D. R. Smith, and A. Moreau, Influence of spatial dispersion in metals on the optical response of deeply subwavelength slit arrays, Phys. Rev. B 93, 045413 (2016).