Third harmonic generation of undoped graphene in Hartree-Fock approximation

J. L. Cheng,[1,2,*] J. E. Sipe,[3,†] and Chunlei Guo[1,4,‡]

[1]*The Guo China-US Photonics Laboratory, Changchun Institute of Optics, fine Mechanics and Physics,*
Chinese Academy of Sciences, 3888 Eastern South Lake Road, Changchun, Jilin 130033, China
[2]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*
[3]*Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 1A7, Canada*
[4]*The Institute of Optics, University of Rochester, Rochester, New York 14627, USA*

We theoretically investigate the effects of Coulomb interaction, at the level of the unscreened Hartree-Fock approximation in an equation of motion framework, on third harmonic generation from undoped graphene. The unperturbed electronic states are described by a widely used two-band tight-binding model and the Coulomb interaction is described by the Ohno potential. The ground state is renormalized by taking into account the Hartree-Fock term and the optical conductivities are obtained by numerically solving the equations of motion. The absolute values of conductivity for third-harmonic generation depend on the photon frequency Ω as Ω^{-n} for $\hbar\Omega < 1$, and then show a peak as $3\hbar\Omega$ approaches the renormalized energy of the M point. Taking into account the Coulomb interaction, n is found to be 5.5, which is significantly greater than the value of 4 found with the neglect of the Coulomb interaction. Therefore, the Coulomb interaction enhances third-harmonic generation at low photon energies—for our parameters $\hbar\Omega < 0.8$ eV—and then reduces it until the photon energy reaches about 2.1 eV. The effect of the background dielectric constant is also considered.

I. INTRODUCTION

The Coulomb interaction between carriers plays an important role in determining the band structure of a crystal and its optical response [1]. In a gapped semiconductor, the repulsive Coulomb interaction leads to the so-called GW correction, which increases the band gap above that obtained from the independent particle approximation (IPA) [2]. In contrast, the attractive Coulomb interaction, usually between electrons and holes, leads to the formation of excitons. Both effects must be included in a calculation to identify the correct linear optical response near the absorption edge of a semiconductor. Besides these contributions, the Coulomb interaction also leads to scattering and the resulting relaxation and thermalization of carriers, beginning on a timescale of tens of femtoseconds. For usual semiconductors, these Coulomb effects occur in the weak interaction regime.

For two-dimensional graphene, the atomic scale thickness leads to strong quantum confinement and reduces the Coulomb screening [3]. Due to the gapless linear band-structure characteristic of massless Dirac fermions, the strength of the Coulomb interaction in graphene, described by the ratio of the Coulomb energy to the kinetic energy [4], is $\alpha_g = e^2/(4\pi\epsilon_0\epsilon\hbar v_F)$, assuming an effective background dielectric constant ϵ and a Fermi velocity v_F. Taking the experimental value $v_F = 10^6$ m/s, the resulting ratio $\alpha_g \approx 2.2/\epsilon$ indicates the Coulomb interaction can be tuned from the weak interaction regime ($\epsilon \approx 37.5$ in certain liquid environments [5]) to the strong interaction regime ($\epsilon = 1$ for a free-standing sample). The Coulomb interaction in graphene affects its optical response in unexpected ways. First, *ab initio* calculations [6] show that the GW correction cannot open the gap, but does increase the Fermi velocity, and corrects the linear dispersion around the Dirac points with a logarithmic function [7]. Second, bound excitons do not exist, and excitonic effects (EXEs) are not important around the Dirac points; thus, the linear optical absorption at low photon energies is not affected by the Coulomb interaction [6,8]. Yet saddle point excitons can be formed around the M point of the band structure, and the corresponding resonant optical absorption is redshifted by about 0.6 eV [5,6], with a Fano-type line shape due to resonance with the continuum electron-hole states [9,10]. More generally, the optical absorption from the infrared to the visible is found to be reduced due to the 2D Coulomb interaction, which also provides a very fast relaxation of hot electrons [11].

While there have been a number of investigations into the effects of the Coulomb interaction on the linear optical response of graphene [5,6,9,10], there have been few theoretical studies that take into account the effects of the Coulomb interaction on the nonlinear optical response [12]. The experimental investigations on second- [13] and third-order optical nonlinearities [14–16], as well as studies of high harmonic generation [17–19], show many advantages of utilizing 2D materials in nonlinear optics [20–22]. These include the extremely large nonlinear coefficients [23,24], the ease of integration in photonic devices [15,25–27], and the tunability by the chemical potential [14–16]. Some of these features can be well predicted and understood from calculations within

---

the IPA [23,28–32]. Because the Coulomb interaction hardly affects the linear optical absorption, it might be natural to assume that its effects on the nonlinear optical response would also be small. However, a recent study by Avetissian and Mkrtchian [12] reported a large enhancement of harmonic generation at THz frequencies due to the 2D Coulomb interaction. There have also been discussions of the plasmonic effects [33–36] on the optical nonlinearity of graphene. In the present paper, we model the Coulomb interaction by the Ohno potential and consider its effects on third-harmonic generation, in the unscreened Hartree-Fock (HF) approximation, for fundamental photon energies between 0.2 eV and 3.5 eV. The effects of a background dielectric constant are also considered. In principle, our approach is appropriate for intrinsic graphene, where the screening should not be important for optical nonlinearity, and also for lightly doped graphene. For the latter, used in most early experiments [37–40], the screening induced by the free carriers is not important for large incident photon energy due to energy mismatch; moreover, the independent particle theory [28] shows that in this case, the third-harmonic generation is not sensitive to a nonvanishing but small chemical potential. For heavily doped graphene, as used in recent experiments [14,16] to manifest the chemical potential related resonances in third harmonic geneartion, the screening of the Coulomb interaction should be taken into account in calculating the optical nonlinearity of graphene. However, due to the complexity both in theory and in numerical calculation, in the present paper we first perform a preliminary study under the unscreened Hartree-Fock approximation, leaving the treatment of screening effects to a future work. And to avoid the effects of screening induced by the optically excited carriers, we limit our calculation to the weak electric field regime.

We organize this paper as the following. In Sec. II, we set up the equation of motion in the Hartree-Fock approximation. In Sec. III, we present our numerical scheme and numerical results for the band structure, the density of states (DOS), the linear conductivity, and the nonlinear conductivity for third-harmonic generation; the effect of the background dielectric constant is also investigated. In Sec. IV, we conclude.

II. MODEL

We describe the dynamics of the electrons by a density matrix with components $\rho_{\alpha\beta\boldsymbol{k}}(t)$, where α and β label the atom sites A or B, and $\boldsymbol{k}$ is a crystal wave vector. With the application of an electric field $\boldsymbol{E}(t)$, the density-matrix components satisfy the equation of motion:

$$i\hbar\partial_t\rho_{\alpha\beta\boldsymbol{k}}(t) = \left[H^0_{\boldsymbol{k}+e\boldsymbol{A}(t)/\hbar} + H^{\mathrm{HF}}_{\boldsymbol{k}}(t), \rho_{\boldsymbol{k}}(t)\right]_{\alpha\beta}$$
$$+ e\boldsymbol{E}(t)\cdot(\boldsymbol{\tau}_\alpha - \boldsymbol{\tau}_\beta)\rho_{\alpha\beta\boldsymbol{k}}(t)$$
$$- i\Gamma\left[\rho_{\alpha\beta\boldsymbol{k}}(t) - \rho^0_{\alpha\beta(\boldsymbol{k}+e\boldsymbol{A}(t)/\hbar)}\right]. \quad (1)$$

Here $H^0_{\boldsymbol{k}}$ is a tight-binding Hamiltonian [28], formed by the p_z orbitals of carbon atoms,

$$H^0_{\alpha\beta\boldsymbol{k}} = -\gamma_0(f_{\boldsymbol{k}}\delta_{\alpha,A}\delta_{\beta,B} + f^*_{\boldsymbol{k}}\delta_{\alpha,B}\delta_{\beta,A}), \quad (2)$$

where γ_0 is the nearest-neighbor hopping parameter and $f_{\boldsymbol{k}} = 1 + e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2}$ is the structure factor, with $\boldsymbol{a}_i$ the primitive lattice vectors and $\boldsymbol{\tau}_\alpha$ the displacement of the αth atom in the

unit cell. The vector potential $\boldsymbol{A}(t)$ satisfies $\partial_t\boldsymbol{A}(t) = -\boldsymbol{E}(t)$. The last term in Eq. (1) describes the relaxation processes phenomenologically, with a relaxation parameter Γ; the system relaxes to an equilibrium state $\rho^0_{\boldsymbol{k}}$ in the moving frame, as will be discussed in detail below. In the Appendix, we give a brief derivation of Eq. (1) based on the tight-binding model.

The carrier-carrier Coulomb interaction is included in the unscreened HF approximation through the term $H^{\mathrm{HF}}_{\alpha\beta\boldsymbol{k}}(t)$, which is given by

$$H^{\mathrm{HF}}_{\alpha\beta\boldsymbol{k}}(t) = -\int\frac{d\boldsymbol{k}'}{(2\pi)^2}V_{\alpha\beta(\boldsymbol{k}-\boldsymbol{k}')}\rho_{\alpha\beta\boldsymbol{k}'}(t), \quad (3)$$

The Coulomb interaction term $V_{\alpha\beta\boldsymbol{k}} = \mathcal{A}\sum_j e^{-i\boldsymbol{k}\cdot\boldsymbol{R}_j}V_{j;\alpha\beta}$ with $\mathcal{A}$ being the area of the unit cell is taken to be the Fourier transform of the Ohno potential [41],

$$V_{i,\alpha\beta} = \frac{U}{\epsilon\sqrt{\left(\frac{4\pi\epsilon_0|\boldsymbol{R}_i+\boldsymbol{\tau}_\alpha-\boldsymbol{\tau}_\beta|U}{e^2}\right)^2+1}}, \quad (4)$$

where U is an onsite energy, ϵ can be considered as an effective background dielectric constant [42], and the $\boldsymbol{R}_i$ are lattice vectors. We separate the HF term into two contributions,

$$H^{\mathrm{HF}}_{\alpha\beta\boldsymbol{k}}(t) = \lambda_m H^{\mathrm{HF};(0)}_{\alpha\beta(\boldsymbol{k}+e\boldsymbol{A}(t)/\hbar)} + \lambda_e H^{\mathrm{HF};(1)}_{\alpha\beta\boldsymbol{k}}(t), \quad (5)$$

$$H^{\mathrm{HF};(0)}_{\alpha\beta\boldsymbol{k}} = -\int\frac{d\boldsymbol{k}'}{(2\pi)^2}V_{\alpha\beta(\boldsymbol{k}-\boldsymbol{k}')}\rho^0_{\alpha\beta\boldsymbol{k}'}, \quad (6)$$

with two auxiliary parameters $\lambda_m = \lambda_e = 1$. The first term exists even in the absence of external field. It takes into account the Coulomb interaction at the level of a mean-field approximation (MFA), and it can modify the band structure significantly [6,43]. Accordingly, it also affects the equilibrium distribution. The second term, $H^{\mathrm{HF};(1)}_{\alpha\beta\boldsymbol{k}}(t) = H^{\mathrm{HF}}_{\alpha\beta\boldsymbol{k}}(t) - H^{\mathrm{HF};(0)}_{\alpha\beta(\boldsymbol{k}+e\boldsymbol{A}(t)/\hbar)}$, describes the interaction between optically excited carriers. To better understand the two contributions, $\lambda_{m/e}$ will be intentionally changed in our numerical calculation to include (= 1) or exclude (= 0) an effect.

Before we can solve Eq. (1), it is important to determine $\rho^0_{\boldsymbol{k}}$, the ground state. While the possibility of forming a new excitonic ground state in the strong interaction regime has been extensively discussed [4,44,45], for most experimental scenarios there is a SiO$_2$ or Si substrate, with a dielectric constant larger than 2; thus the Coulomb interaction should be in the weak interaction regime. Therefore, we limit ourselves to a ground state within the IPA. In our treatment, the Coulomb interaction has two main consequences: First, it modifies the band structure through $H^{\mathrm{HF};(0)}_{\alpha\beta\boldsymbol{k}}$. The eigenenergies $\varepsilon_{s\boldsymbol{k}}$ and eigenstates $C_{s\boldsymbol{k}} = \begin{pmatrix}C^A_{s\boldsymbol{k}}\\C^B_{s\boldsymbol{k}}\end{pmatrix}$ with a band index $s = \pm$ are determined from the Schrödinger equation,

$$\sum_\beta\left(H^0_{\alpha\beta\boldsymbol{k}} + H^{\mathrm{HF};(0)}_{\alpha\beta\boldsymbol{k}}\right)C^\beta_{s\boldsymbol{k}} = \varepsilon_{s\boldsymbol{k}}C^\alpha_{s\boldsymbol{k}}, \quad (7)$$

and the equilibrium density matrix is calculated from

$$\rho^0_{\alpha\beta\boldsymbol{k}} = \sum_s\frac{C^\alpha_{s\boldsymbol{k}}(C^\beta_{s\boldsymbol{k}})^*}{1+e^{(\varepsilon_{s\boldsymbol{k}}-\mu)/(k_BT)}} \quad (8)$$

for a specified chemical potential μ and temperature T. Equations (6)–(8) form a self-consistent set of equations, and

they are solved iterately. By initially setting $H_{\alpha\beta k}^{\mathrm{HF};(0)} = 0$ in Eq. (7), we get C_{sk}^{β}; next we calculate ρ_k^0 using Eq. (8), and then we update $H_{\alpha\beta k}^{\mathrm{HF};(0)}$ using Eq. (6), repeating this procedure until $\rho_{\alpha\beta k}^0$ is converged. A second consequence of the Coulomb interaction is that the term $H_{\alpha\beta k}^{\mathrm{HF};(1)}(t)$ leads to the excitonic effects in the framework of carrier dynamics [46].

In this paper, the main quantity of interest is the current density, which is given by $\boldsymbol{J}(t) = (2\pi)^{-2} \int d\boldsymbol{k} \boldsymbol{J}_k(t)$ with

$$\boldsymbol{J}_k(t) = -e \sum_{\alpha\beta} \boldsymbol{v}_{\beta\alpha k}(t)\rho_{\alpha\beta k}(t), \qquad (9)$$

$$\boldsymbol{v}_{\alpha\beta k}(t) = \frac{1}{i\hbar}\left\{(\boldsymbol{\tau}_\alpha - \boldsymbol{\tau}_\beta + i\boldsymbol{\nabla}_k)\left[H_{\alpha\beta(k+eA(t)/\hbar)}^0 + H_{\alpha\beta k}^{\mathrm{HF}}(t)\right]\right\}. \qquad (10)$$

III. RESULTS

In our numerical calculation, we adopt the parameters $\gamma_0 = 2.34$ eV, $U = 8.29$ eV [47], $\hbar/\Gamma = 20$ fs [30], $\mu = 0$ eV, $T = 300$ K, and divide the Brillouin zone into a $N \times N$ grid with $N = 1500$; our main conclusions are not very sensitive to the exact values of γ_0 and U. The time differential is discretized by a fourth-order Runge-Kutta method with a time step 0.05 fs. We consider background dielectric constants ϵ varying between 2 and 9. The results are presented for three different approaches: the IPA (with $\lambda_m = \lambda_e = 0$), the MFA (with $\lambda_m = 1$ and $\lambda_e = 0$), and taking into account EXEs (with $\lambda_m = \lambda_e = 1$).

We numerically calculate the current density for a periodic field $\boldsymbol{E}(t) = E_0\hat{\boldsymbol{x}}e^{-i\Omega t} + $ c.c., with $E_0 = 5\times10^6$ V/m, in the time range $0 \leqslant t \leqslant 300$ fs [48]. Because the relaxation time is much shorter than 300 fs, we can approximate the current density in the last period as $J_x(t) = \sum_n [J_x^{(n)}e^{-in\Omega t} + $ c.c.]. The linear conductivity can be obtained by $\sigma^{(1)}(\Omega) = J_x^{(1)}/E_0$ and the THG conductivity is obtained by $\sigma^{(3)}(\Omega) = J_x^{(3)}/E_0^3$. For THG, the photon energies are chosen in the range of $0.2 \leqslant \hbar\Omega \leqslant 3.4$ eV. The nonlinear response at terahertz frequencies is also very interesting, of course. However, unlike the numerical scheme used by Avetissian and Mkrtchian [12], the application of our approach in the terahertz regime would require a very dense grid and a very long evolution time. As well, our neglect of the dynamic screening would likely

affect the reasonableness of the results for terahertz response. Therefore, we leave the treatment of terahertz response to a future study.

Insight into the dynamics induced by the applied field can be gained by looking at the transition-energy-resolved conductivities, which are defined as

$$\sigma^{(n)}(\Omega, \epsilon_t) = \int \frac{d\boldsymbol{k}}{(2\pi)^2} \frac{J_{x:\boldsymbol{k}}^{(n)}}{E_0^n} \delta[\epsilon_t - (\varepsilon_{+\boldsymbol{k}} - \varepsilon_{-\boldsymbol{k}})], \qquad (11)$$

where $J_{x:\boldsymbol{k}}^{(n)}$ is the nth-order Fourier transformation of $J_{x:\boldsymbol{k}}(t)$. They describe how the electronic states at a given transition energy ϵ_t contribute to the optical conductivity at a fundamental frequency Ω. In the IPA and MFA approximations, $|\sigma^{(n)}(\Omega, \epsilon_t)|$ exhibits a peak at $\epsilon_t = m\hbar\Omega$ for the m-photon resonant optical transition for $1 \leqslant m \leqslant n$, as would be expected within a single-particle description. In Fig. 1(a), we illustrate the electronic states for m-photon resonant transitions at $m = 1, 2, 3$. With the inclusion of the HF term, the energy shift of these peaks can be used to identify EXEs. Very generally, the conductivities are given by the integral over transition energies of these resolved quantities, $\sigma^{(n)}(\Omega) = \int d\epsilon_t \sigma^{(n)}(\Omega, \epsilon_t)$.

A. Band structure and linear conductivity

Figures 1(a) and 1(b) give the single particle band structure and DOS in the IPA and the MFA for $\epsilon = 3$. In the IPA, the dispersion is linear around the Dirac points with a Fermi velocity $v_F = \sqrt{3}a_0\gamma_0/2\hbar \approx 0.76\times10^6$ m/s, and the DOS is also approximately linear in a large energy range between about -1.5 eV and 1.5 eV. There are significant changes in the band structure as we move to the MFA. Our results are in line with a host of other calculations, showing that the single particle bands deviate from linear dispersion [43], and even away from the Dirac points, the Fermi velocity increases. The band energy at the M point increases from 2.34 eV in the IPA to 3.25 eV in the MFA, and the energy at the Γ point also increases from 7 eV in the IPA to 8.27 eV in the MFA. The increase of these characteristic energies by the HF term shows behavior similar to the effects of GW corrections in gapped semiconductors. Although both the IPA and MFA are single-particle approximations, their different band structures result in a clear difference between their predicted
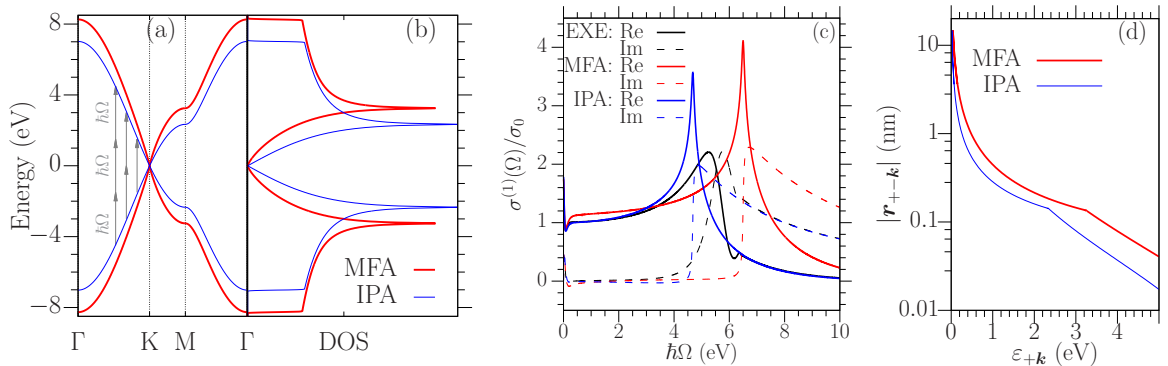

FIG. 1. Effects of the HF term on (a) the band structure, (b) the density of states, (c) the linear optical conductivity, and (d) the energy dependence of the dipole transition matrix elements between the two bands along the high symmetry lines K → M → Γ. The gray arrows in (a) indicate the resonant transitions by multiple photons.
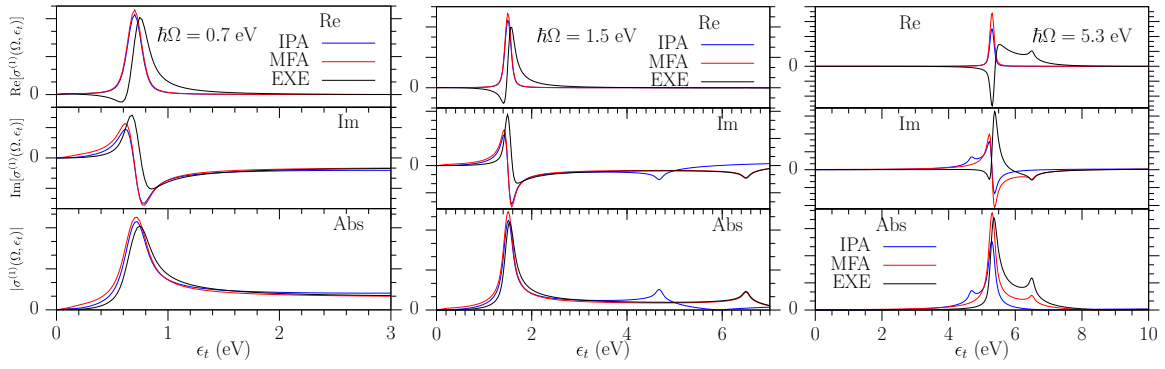
FIG. 2. Transition-energy-resolved linear conductivity for different excitation frequencies $\hbar\Omega = 0.7$, 1.5, and 5.3 eV in the IPA, MFA, and EXE models. The values for the real part, the imaginary part, and the absolute values are presented.

linear optical conductivities, as shown in Fig. 1(c). The MFA corrections to the dispersion also enhance the real parts of the linear conductivity for photon energies less than 2.5 eV. The enhancement is a joint effect involving the enhanced dipole matrix elements shown in Fig. 1(d) and the decreased DOS shown in Fig. 1(b), with the former dominating. The result shows how the MFA modifies the single-particle electronic states. When the photon energy matches the optical transition energy at the M point, the real part of the conductivity shows a very sharp peak in both the IPA and MFA, induced by the Van Hove singularity at the M point. With the further inclusion of EXEs, the resulting linear conductivity in the EXE exhibits three main features: (1) The conductivity for $\hbar\omega < 3$ eV is very close to the universal conductivity obtained in the IPA. (2) The singularity peak is shifted from a photon energy 6.5 eV to 5.3 eV, which indicates a strong EXE around the M point. The exciton binding energy is about 1.2 eV, of the same order of magnitude for excitons in other 2D materials. (3) The broadened peak at the M point, followed by a small dip, confirms the Fano resonance between the saddle excitons and continuum electron-hole states. These results agree qualitatively with an *ab initio* calculation [6], which provides a good check of the reasonableness of our model. Roughly speaking, the linearity conductivity plot for EXE lies between that of IPA and MFA, indicating an interference between the mean-field and excitonic contributions. We will see this kind of "undoing" of the mean-field contributions by the excitonic contributions is even more pronounced when we consider the nonlinear response below.

In Fig. 2, we plot the transition-energy-resolved conductivity, $\sigma^{(1)}(\Omega, \epsilon_t)$, for the excitation photon energies $\hbar\Omega = 0.7$, 1.5, and 5.3 eV. For the single-particle models (the IPA and MFA), a Lorentzian shape $\sigma^{(1)}(\Omega, \epsilon_t) \propto i(\hbar\Omega - \epsilon_t + i\Gamma)^{-1}$ results. The real part of $\sigma^{(1)}(\Omega, \epsilon_t)$, which corresponds to the absorption at each transition energy ϵ_t, is always positive and localized around $\epsilon_t \sim \hbar\Omega$ with a broadening approximately determined by the relaxation parameter; this is as expected for a single-particle theory. The difference between the IPA and the MFA is mainly due to the different optical transition matrix elements. For the imaginary part, the band structure and the DOS play important roles, with peaks around the singularity at the M points ($\epsilon_t \approx 4.6$ eV for the IPA and 6.5 eV for the MFA). With the inclusion of the EXEs, $\sigma^{(1)}(\Omega, \epsilon_t)$ deviates from its MFA behavior only around the peak at

$\epsilon_t \approx \hbar\Omega$. For $\hbar\Omega = 0.7$ and 1.5 eV, the absolute value only changes slightly, but the mixture of the real and imaginary parts indicates a phase change of the optical transition matrix elements. Overall, EXEs are very weak for low-photon energies, agreeing with the *ab initio* results. For $\hbar\Omega = 5.3$ eV, the absolute value of $\sigma^{(1)}(\Omega, \epsilon_t)$ includes a significant contribution from the electronic states at the M point for $\epsilon_t > \hbar\Omega$, showing the important role played by EXEs.

B. Third-harmonic generation

We now turn to THG, the conductivities of which are plotted in Fig. 3 for IPA, MFA, and EXEs. We first look at the spectra in the IPA. The absolute value of conductivity decreases with the photon energy, approximately following a power law $\propto \Omega^{-n_0}$ with $n_0 \sim 4.4$, and reaches a minimum at $\hbar\Omega \approx 1.2$ eV; then it increases to a maximum at $\hbar\Omega \sim 1.6$ eV, and decreases again afterward. This spectrum is very similar to perturbative results in literature [28,49]. Ignoring relaxation, the analytic perturbative conductivity [29] gives $\sigma_{\text{THG}} \propto \Omega^{-4}$ for low photon energies, shown in Fig. 3 as
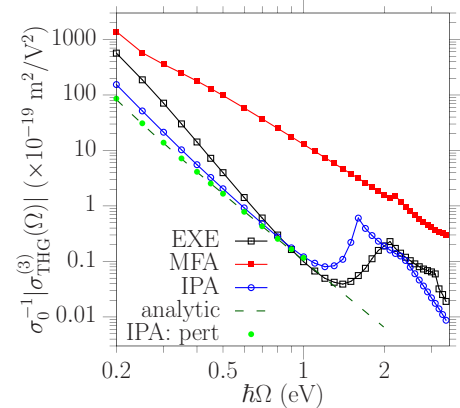

FIG. 3. Effects of the HF term on the nonlinear conductivity for third-harmonic generation. The dotted curve gives the perturbative conductivity obtained by an analytic expression [29]; the other three curves are obtained under a field $E_0 = 5 \times 10^6$ V/m. The blue (red) curves are the results obtained from the single-particle band structure with (without) HF corrections, and the black curves are obtained with the excitonic effects. The green dots are obtained using a field $E_0 = 10^6$ V/m.
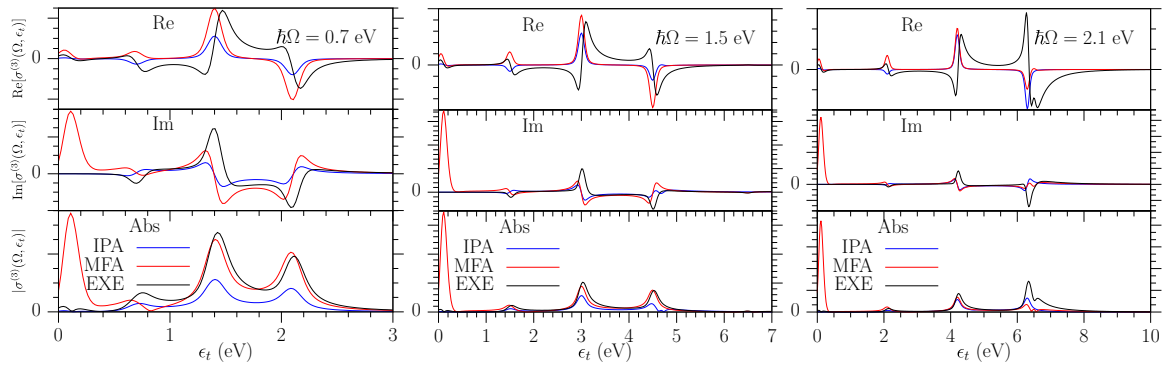
FIG. 4. Transition-energy-resolved THG conductivity for fundamental photon energies $\hbar\Omega = 0.7$ eV, 1.5 eV, and 2.1 eV.

a dashed curve. Our numerical results give a faster decay ($n_0 > 4$) because the simulation field strength $E_0 = 5\times10^6$ V/m is slightly beyond the perturbative limit. Therefore, saturation effects start to play a role, and they enhance the THG due to the extra optically excited carriers from the one-photon absorption [30]. As might be expected, the saturation is more important at low-photon energy, where the saturation intensity is lower. The perturbative limit is obtained for a weaker field $E_0 = 10^6$ V/m, as shown by the green filled dots in Fig. 3, which agree with the analytic results very well. Because a much longer simulation time for this weak field is required, we keep our other calculations using $E_0 = 5\times10^6$ V/m. The peak around 1.6 eV is induced by the three-photon resonant transition at the Van Hove singularity point. This is consistent with an analytic perturbative result [49].

When the Coulomb interaction is included at the level of the MFA, the spectra are obviously different, shown as filled red squares in Fig. 3. For the calculated photon energies, the values in the MFA are orders of magnitude larger than those in the IPA, mirroring the change in the linear conductivity as we move from the IPA to the MFA, although the increase is much greater here. Except for a very weak peak around $\hbar\Omega \approx 2.2$ eV, which corresponds to the three-photon resonant transition at the M point, the whole spectrum follows a power law $\propto \Omega^{-n_m}$ with $n_m = 3$ for $\hbar\Omega < 3.4$ eV. When the EXEs are included, the THG conductivity again changes dramatically, as shown in the black square in Fig. 3. At low-photon energies, $\hbar\Omega < 1$ eV, the THG shows a different scaling with Ω than seen in either the IPA or the MFA, scaling as Ω^{-n_e} with $n_e \sim 5.5$; compared to the IPA results, the EXE results are enhanced by about three times at $\hbar\Omega = 0.2$ eV, but reduced about 20% at $\hbar\Omega = 1$ eV. When the photon energy increases, a peak appears at $\hbar\omega \approx 2.1$ eV, which is the same energy for the weak peak appearing in the MFA. In brief, the Coulomb interaction mainly affects the THG in the following ways: (1) THG in the MFA is orders of magnitude larger than that in the IPA. The increase is even much larger than that calculated for the linear response. Therefore, the details of the band structure are very important for understanding the optical nonlinearity. (2) The further inclusion of EXEs can bring the THG very close to the results in the IPA, showing a strong interference between the mean-field contribution and the excitonic contribution. (3) At low-photon energies, the Coulomb interaction changes the photon energy dependence, with the power index changing from $n_0 = 4.4$ in the IPA to

$n_m = 3$ in the MFA and $n_e = 5.5$ in the EXE results. (4) For the EXE spectra, the three-photon resonance with the M-point gives a peak at $\hbar\Omega \approx 2.1$ eV. This energy does not correspond to the one third of the M-point exciton energy (5.3 eV/3 $\sim$ 1.8 eV), but is very close to one-third of the M-point energy in the MFA (6.5 eV/3 $\sim$ 2.16 eV). The lack of an energy shift may indicate that the EXEs are not important for the three-photon resonance with the M point. These four features indicate the importance of the Coulomb interaction on the THG in an intrinsic graphene; both (1) and (2) correspond to similar effects of Coulomb interaction on the linear optical response.

To gain some insight into these features, we calculate the transition-energy-resolved THG conductivity, shown in Fig. 4 for $\hbar\Omega = 0.7$ eV, 1.5 eV, and 2.1 eV. In the single-particle approximations, the values of $|\sigma^{(3)}(\Omega, \epsilon_t)|$ show three peak contributions located around transition energies $m\hbar\Omega$ with $m = 1, 2, 3$, corresponding to the one-photon, two-photon, and three-photon resonant transitions. In the IPA, the two-photon resonant transition destructively interferes with the one-photon and three-photon resonant transitions [28] (approximately corresponding to prefactors $-17, 64, -45$), which results in a small THG conductivity. In the MFA, there are of course changes in the DOS and the dipole matrix elements that affect the THG, but as well the interference is also greatly changed due to the modification of the Dirac band structure. Although there is still some destructive interference between these transitions, the sign of the first peak in the MFA is changed compared to that in the IPA, the cancellation is not complete, and this leads to a larger THG response. As well, there is an additional peak located at zero energy. This peak has the same sign as the peak at $2\hbar\Omega$, and it is larger in the MFA than in the IPA. With the inclusion of excitonic effects, the transition energy distribution of the absolute values of THG becomes very similar to that of the IPA, especially around the peak at $\epsilon_t = 0$; simultaneously, an additional phase is introduced around each peak that changes both the real and imaginary parts. The results indicate that the interference between the mean-field contributions and the EXEs also exists in the optical nonlinearity. For $\hbar\Omega = 2.1$ eV, the transition-energy-resolved spectra shows that the resonant peaks locate around $\epsilon_t \sim m\hbar\Omega$, which are similar to the dependence in the single-particle approximation. This confirms that EXEs play a minor role for the peak around $\hbar\Omega \approx 2.1$ eV in Fig. 3.
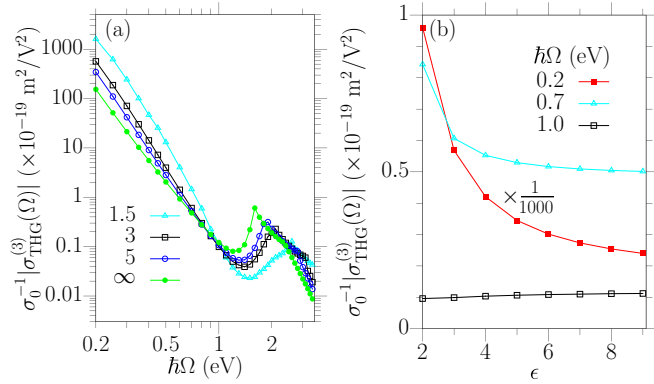
FIG. 5. Effects of the substrate dielectric constant ϵ on the THG conductivity. (a) The spectrum of the THG conductivity for different ϵ. (b) The ϵ of the THG conductivity for three-photon energy $\hbar\Omega = 0.2$, 0.7, and 1.0 eV. The results for 0.2 eV are scaled by 10^{-3}.

We now turn to a comparison with early experiments [37–40], in which lightly doped graphene samples were used. For the experiments with wavelengths [37–39] in the range of $1.5 \sim 1.7\,\mu$m, the calculated THG coefficients in the IPA and EXE are almost the same, and thus the inclusion of the Coulomb interaction at the Hartree-Fock level does not improve the results in the IPA and linear dispersion approximation. For a fundamental wavelength at $\lambda = 0.8\,\mu$m, with its third harmonics resonant with the M point [40], the linear dispersion approximation is no longer appropriate, and our numerical calculation gives the THG susceptibility $\chi^{(3)} = \sigma^{(3)}\lambda/(-6i\pi c\epsilon_0 d_{\text{eff}})$ as 1.8×10^{-19} m^3/V^3 for the peak of the IPA at $\hbar\Omega \sim 1.6$ eV and 0.68×10^{-19} m^3/V^3 for the peak of the EXE at $\hbar\Omega \sim 2.1$ eV, where c is the speed of light and $d_{\text{eff}} = 3.3$ Å is the effective thickness of graphene. Both calculated values are of the same order of magnitude of the experimental value [40].

C. Substrate effects

Because of the complexity of the nonlinear optical response of graphene, it is easy to reveal the consequences of many-body effects by changing the substrate, or more generally the environment, which is very effective in tuning the interaction strength of the Coulomb interaction [5]. In Fig. 5, we show how the background dielectric constant ϵ affects the THG. The Coulomb interaction is inversely proportional to ϵ, and thus a large background dielectric constant corresponds to a weak interaction; the limit $\epsilon \to \infty$ corresponds to no Coulomb interaction.

With an increase in ϵ, the THG conductivity decreases for photon energies $\hbar\Omega$ less than about 0.8 eV, and the power index n for $\sigma^{(3)} \propto \Omega^{-n}$ also decreases. But for photon energies $\hbar\Omega$ larger than about 0.8 eV, the THG conductivity increases; the peak corresponding to the resonant three-photon transition at the M point is shifted to a lower photon energy, because the energy renormalization decreases with the strength of the Coulomb interaction. The THG conductivity is more sensitive to the substrate dielectric constant at lower photon energies, and it can change over one order of magnitude for $\hbar\Omega = 0.2$ eV when ϵ changes from 2 to ∞.

IV. CONCLUSION

We have theoretically investigated the effects of the Coulomb interaction on third-harmonic generation from undoped graphene in the unscreened Hartree-Fock approximation, with the inclusion of mean-field energy corrections and EXEs. Although there are no bound excitons formed in gapless graphene, the Coulomb interaction still affects the third-harmonic generation significantly. We find that the Coulomb interaction can increase the amplitude of third-harmonic generation at low-photon energy, and decrease it at high-photon energy. Despite the fact that saddle point excitons lead to a large energy shift of the resonant peak in linear absorption, they leave no fingerprint on the energy of the three-photon resonance in third harmonic generation. The underlying physics can be understood by the inclusion of the Coulomb interaction in a step-by-step fashion. At the level of the MFA, the Coulomb interaction greatly modifies the single-particle band structure, which leads to an enhancement of the third-harmonic generation by around two orders of magnitude compared to the results obtained without the Hartree-Fock term. However, with the full inclusion of the Hartree-Fock term, this large enhancement is largely canceled, except for the Coulomb-induced changes of the power scaling. Therefore, there is a very strong interference between the contributions from the mean-field energy correction and EXEs. We found these effects could be revealed experimentally by changing the background dielectric constant, leading to changes both in the absolute value of the third-harmonic generation coefficient and in its frequency dependence. The strength of the modifications due to changes in the environment is very sensitive to the fundamental frequency and a stronger dependence on the environment dielectric constant can be found at lower incident photon energy. Therefore, the strong dependence of third-harmonic generation on the Coulomb interaction may provide an tool for studying the many-body effect in graphene.

For undoped graphene, our results show that the Coulomb interaction can significantly modify third-harmonic generation at low frequencies. Thus it may be important as well in other nonlinear optical phenomena involving small frequencies, such as degenerate four-wave mixing, coherent current injection, Kerr effects and two-photon absorption, current-induced second-order optical nonlinearity, and jerk current [50]. For doped graphene, where dynamic screening should be important, the unscreened Hartree-Fock approximation may be not adequate. Previous calculations [28,31] of the optical nonlinearity in doped graphene exhibit many resonances, arising whenever any photon energy involved matches the gap induced by a nonzero chemical potential; they are also observed in recent experiments [14,16]. What role the Coulomb interaction plays around these resonances is an important but unexplored problem in understanding the physics of the optical nonlinearity in graphene.

APPENDIX: THE DERIVATION OF THE EQUATION OF MOTION

We start with a tight-binding Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_c + e\boldsymbol{E}(t) \cdot \hat{\boldsymbol{r}}$ with

$$\hat{H}_0 = \sum_{i\alpha, j\beta, \sigma} \gamma_{i-j,\alpha\beta} a_{i\alpha\sigma}^\dagger a_{j\beta\sigma}, \tag{A1}$$

$$\hat{H}_c = \frac{1}{2} \sum_{\substack{i\alpha\sigma_1 \\ j\beta\sigma_2}} V_{i-j,\alpha\beta} a_{i\alpha\sigma_1}^\dagger a_{j\beta\sigma_2}^\dagger a_{j\beta\sigma_2} a_{i\alpha\sigma_1}, \tag{A2}$$

$$\hat{\boldsymbol{r}} = \sum_{i\alpha\sigma} \boldsymbol{R}_{i\alpha} a_{i\alpha\sigma}^\dagger a_{i\alpha\sigma}. \tag{A3}$$

Here $a_{i\alpha\sigma}$ is an annihilation operator of an electronic p_z orbital with a spin σ ($=\uparrow, \downarrow$) at the site $\boldsymbol{R}_{i\alpha} = \boldsymbol{R}_i + \boldsymbol{\tau}_\alpha$, where $i = (n_i, m_i)$ is an abbreviated notation for $\boldsymbol{R}_i = n_i \boldsymbol{a}_1 + m_i \boldsymbol{a}_2$. The term $\hat{H}_0$ gives the unperturbated tight binding Hamiltonian. In this paper, we use the parameters

$$\gamma_{i,AB} = -\gamma_0 [\delta_{i,(0,0)} + \delta_{i,(1,0)} + \delta_{i,(0,1)}], \tag{A4}$$

$$\gamma_{i,BA} = -\gamma_0 [\delta_{i,(0,0)} + \delta_{i,(-1,0)} + \delta_{i,(0,-1)}]. \tag{A5}$$

We treat the Coulomb interaction $\hat{H}_c$ at the level of Hartree-Fock approximation, and take

$$\hat{H}_c \rightarrow \sum_{\substack{i\alpha\sigma_1 \\ j\beta\sigma_2}} V_{i-j,\alpha\beta} [\langle a_{i\alpha\sigma_1}^\dagger a_{i\alpha\sigma_1} \rangle a_{j\beta\sigma_2}^\dagger a_{j\beta\sigma_2}$$
$$- \langle a_{i\alpha\sigma_1}^\dagger a_{j\beta\sigma_2} \rangle a_{j\beta\sigma_2}^\dagger a_{i\alpha\sigma_1}]. \tag{A6}$$

The constant energy shift in this approximation has been ignored. The notation $\langle P \rangle$ stands for the statistical average of the operator P over the ground state. Considering a paramagnetic ground state in extended graphene, the term $\langle a_{i\alpha\sigma_1}^\dagger a_{j\beta\sigma_2} \rangle$ should only be a function of $i - j$ due to the translational symmetry. Then the first term can be seen to be a simple energy shift. By ignoring the spin we get an effective Hamiltonian as

$$\hat{H}_{\text{eff}}(t) = \sum_{i\alpha j\beta} [\gamma_{i-j,\alpha\beta} + \overline{H}_{i-j,\alpha\beta}^{\text{HF}}] a_{i\alpha}^\dagger a_{j\beta} + e\boldsymbol{E}(t) \cdot \hat{\boldsymbol{r}}, \tag{A7}$$

with

$$\overline{H}_{i-j,\alpha\beta}^{\text{HF}} = -V_{i-j,\alpha\beta} \langle a_{j\beta}^\dagger a_{i\alpha} \rangle. \tag{A8}$$

In this work, we describe the dynamics of the system by a density matrix $\overline{\rho}_{i-j,\alpha\beta}(t) = \langle a_{j\beta}^\dagger(t) a_{i\alpha}(t) \rangle$, with $a_{i\alpha}(t)$ being the operator $a_{i\alpha}$ in Heisenberg representation. In the Hartree-Fock approximation, it satisfies the equation of motion

$$i\hbar \partial_t \overline{\rho}_{i;\alpha\beta}(t) = \sum_j [\gamma_{i-j} + \overline{H}_{i-j}^{\text{HF}}, \overline{\rho}_j(t)]_{\alpha\beta}$$
$$+ e\boldsymbol{E}(t) \cdot (\boldsymbol{R}_i + \boldsymbol{\tau}_\alpha - \boldsymbol{\tau}_\beta) \overline{\rho}_{i;\alpha\beta}(t)$$
$$- i\Gamma [\overline{\rho}_{i;\alpha\beta}(t) - \overline{\rho}_{i;\alpha\beta}^0]. \tag{A9}$$

In the commutator, γ_i, H_i^{HF}, and $\overline{\rho}_i(t)$ are treated as matrices with elements determined by the indexes $\alpha\beta$. The last term models the relaxation process with one phenomenological energy parameter Γ, and $\rho_{i;\alpha\beta}^0$ gives the equilibrium distribution including the Hartree-Fock effects. The velocity operator is given by $\hat{\boldsymbol{v}} = (i\hbar)^{-1}[\hat{\boldsymbol{r}}, \hat{H}_{\text{eff}}(t)] = \sum_{ij\alpha\beta} \boldsymbol{v}_{i-j,\alpha\beta} a_{i\alpha}^\dagger a_{j\beta}$ with

$$\boldsymbol{v}_{j,\alpha\beta} = (i\hbar)^{-1}(\boldsymbol{R}_j + \boldsymbol{\tau}_\alpha - \boldsymbol{\tau}_\beta)(\gamma_{j,\alpha\beta} + \overline{H}_{j;\alpha\beta}^{\text{HF}}). \tag{A10}$$

Then the current density can be calculated through

$$\boldsymbol{J}(t) = -e\mathcal{A}^{-1} \sum_{j\alpha\beta} \boldsymbol{v}_{j,\alpha\beta} \rho_{-j;\beta\alpha}(t). \tag{A11}$$

After Fourier transform of the index i to wave vectors $\boldsymbol{k}$, i.e., $\gamma_j \rightarrow H_{\boldsymbol{k}}^0 = \sum_j e^{-i\boldsymbol{k} \cdot \boldsymbol{R}_j} \gamma_j$, as well as $\overline{H}_j^{\text{HF}} \rightarrow H_{\boldsymbol{k}}^{\text{HF}}$, $\overline{\rho}_j \rightarrow \rho_{\boldsymbol{k}}$, and changing to the moving frame, we get the equations in the main text.

[1] M. Kira and S. Koch, Prog. Quantum Electron. **30**, 155 (2006).

[2] S. G. Louie, Predicting materials and properties: Theory of the ground and excited state, in *Conceptual Foundations of Materials—A Standard Model for Ground- and Excited-State Properties* (Elsevier, Amsterdam, 2006), Chap. 2, pp. 9–53.

[3] E. H. Hwang and S. Das Sarma, Phys. Rev. B **77**, 081412(R) (2008).

[4] J. Grönqvist, T. Stroucken, M. Lindberg, and S. Koch, Eur. Phys. J. B **85**, 395 (2012).

[5] P. Yadav, P. K. Srivastava, and S. Ghosh, Nanoscale **7**, 18015 (2015).

[6] L. Yang, J. Deslippe, C.-H. Park, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **103**, 186802 (2009).

[7] J. Jung and A. H. MacDonald, Phys. Rev. B **84**, 085446 (2011).

[8] M. I. Katsnelson, Europhys. Lett. **84**, 37001 (2008).

[9] K. F. Mak, J. Shan, and T. F. Heinz, Phys. Rev. Lett. **106**, 046401 (2011).

[10] K. F. Mak, F. H. da Jornada, K. He, J. Deslippe, N. Petrone, J. Hone, J. Shan, S. G. Louie, and T. F. Heinz, Phys. Rev. Lett. **112**, 207401 (2014).

[11] E. Malic, T. Winzer, E. Bobkin, and A. Knorr, Phys. Rev. B **84**, 205406 (2011).

[12] H. K. Avetissian and G. F. Mkrtchian, Phys. Rev. B **97**, 115454 (2018).

[13] Y. Zhang, D. Huang, Y. Shan, T. Jiang, Z. Zhang, K. Liu, L. Shi, J. Cheng, J. E. Sipe, W.-T. Liu, and S. Wu, Phys. Rev. Lett. **122**, 047401 (2019).

[14] T. Jiang, D. Huang, J. Cheng, X. Fan, Z. Zhang, Y. Shan, Y. Yi, Y. Dai, L. Shi, K. Liu, C. Zeng, J. Zi, J. E. Sipe, Y.-R. Shen, W.-T. Liu, and S. Wu, Nat. Photon. **12**, 430 (2018), and references therein.

[15] K. Alexander, N. A. Savostianova, S. A. Mikhailov, B. Kuyken, and D. V. Thourhout, ACS Photon. **4**, 3039 (2017).

[16] G. Soavi, G. Wang, H. Rostami, D. G. Purdie, D. De Fazio, T. Ma, B. Luo, J. Wang, A. K. Ott, D. Yoon, S. A. Bourelle,

J. E. Muench, I. Goykhman, S. Dal Conte, M. Celebrano, A. Tomadin, M. Polini, G. Cerullo, and A. C. Ferrari, Nat. Nanotechnol. **13**, 583 (2018).

[17] N. Yoshikawa, T. Tamaya, and K. Tanaka, Science **356**, 736 (2017).

[18] I. Al-Naib, J. E. Sipe, and M. M. Dignam, Phys. Rev. B **90**, 245423 (2014).

[19] M. Baudisch, A. Marini, J. D. Cox, T. Zhu, F. Silva, S. Teichmann, M. Massicotte, F. Koppens, L. S. Levitov, F. J. G. de Abajo, and J. Biegert, Nat. Commun. **9**, 1018 (2018).

[20] F. Bonaccorso, Z. Sun, T. Hasan, and A. C. Ferrari, Nat. Photon. **4**, 611 (2010).

[21] Z. Sun, A. Martinez, and F. Wang, Nat. Photon. **10**, 227 (2016).

[22] A. Autere, H. Jussila, Y. Dai, Y. Wang, H. Lipsanen, and Z. Sun, Adv. Mater. **30**, 1705963 (2018).

[23] S. A. Mikhailov, Europhys. Lett. **79**, 27002 (2007).

[24] E. Hendry, P. J. Hale, J. Moger, A. K. Savchenko, and S. A. Mikhailov, Phys. Rev. Lett. **105**, 097401 (2010).

[25] T. Gu, N. Petrone, J. F. McMillan, A. van der Zande, M. Yu, G. Q. Lo, D. L. Kwong, J. Hone, and C. W. Wong, Nat. Photon. **6**, 554 (2012).

[26] N. Vermeulen, D. Castelló-Lurbe, J. L. Cheng, I. Pasternak, A. Krajewska, T. Ciuk, W. Strupinski, H. Thienpont, and J. Van Erps, Phys. Rev. Appl. **6**, 044006 (2016).

[27] N. Vermeulen, D. Castelló-Lurbe, M. Khoder, I. Pasternak, A. Krajewska, T. Ciuk, W. Strupinski, J. Cheng, H. Thienpont, and J. Van Erps, Nat. Commun. **9**, 2675 (2018).

[28] J. L. Cheng, N. Vermeulen, and J. E. Sipe, New J. Phys. **16**, 053014 (2014); **18**, 029501 (2016).

[29] J. L. Cheng, N. Vermeulen, and J. E. Sipe, Phys. Rev. B **91**, 235320 (2015); **93**, 039904(E) (2016).

[30] J. L. Cheng, N. Vermeulen, and J. E. Sipe, Phys. Rev. B **92**, 235307 (2015).

[31] S. A. Mikhailov, Phys. Rev. B **93**, 085403 (2016).

[32] H. Rostami and M. Polini, Phys. Rev. B **93**, 161411(R) (2016).

[33] S. A. Mikhailov, Phys. Rev. B **84**, 045432 (2011).

[34] H. Rostami, M. I. Katsnelson, and M. Polini, Phys. Rev. B **95**, 035416 (2017).

[35] T. J. Constant, S. M. Hornett, D. E. Chang, and E. Hendry, Nat. Phys. **12**, 124 (2016).

[36] J. L. Cheng, J. E. Sipe, N. Vermeulen, and C. Guo, J. Phys. Photon. **1**, 015002 (2019).

[37] N. Kumar, J. Kumar, C. Gerstenkorn, R. Wang, H.-Y. Chiu, A. L. Smirl, and H. Zhao, Phys. Rev. B **87**, 121406(R) (2013).

[38] A. Säynätjoki, L. Karvonen, J. Riikonen, W. Kim, S. Mehravar, R. A. Norwood, N. Peyghambarian, H. Lipsanen, and K. Kieu, ACS Nano **7**, 8441 (2013).

[39] R. I. Woodward, R. T. Murray, C. F. Phelan, R. E. P. de Oliveira, T. H. Runcorn, E. J. R. Kelleher, S. Li, E. C. de Oliveira, G. J. M. Fechine, G. Eda, and C. J. S. de Matos, 2D Mater. **4**, 011006 (2017).

[40] S.-Y. Hong, J. I. Dadap, N. Petrone, P.-C. Yeh, J. Hone, and R. M. Osgood, Jr., Phys. Rev. X **3**, 021014 (2013).

[41] J. Jiang, R. Saito, G. G. Samsonidze, A. Jorio, S. G. Chou, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **75**, 035407 (2007).

[42] For graphene embedded inside two different materials, the background dielectric constant is the average of the dielectric constants. Because the interaction is mainly through the electric force outside of the graphene plane, the screening from the graphene electrons is ignored.

[43] E. H. Hwang, Ben Yu-Kuang Hu, and S. Das Sarma, Phys. Rev. Lett. **99**, 226801 (2007).

[44] T. Stroucken, J. H. Grönqvist, and S. W. Koch, J. Opt. Soc. Am. B **29**, A86 (2012).

[45] T. Stroucken, J. H. Grönqvist, and S. W. Koch, Phys. Rev. B **84**, 205445 (2011).

[46] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer, Berlin, 2007).

[47] Z. S. Sadeq, R. A. Muniz, and J. E. Sipe, Phys. Rev. Mater. **2**, 014001 (2018).

[48] This work considers the third harmonic generation in the weak field regime, thus the use of a periodic field follows the standard perturbative treatment, which should be suitable in experiments where the pulse duration is much longer than the relaxation time [30].

[49] Z. Liu, C. Zhang, and J. C. Cao, Phys. Rev. B **96**, 035206 (2017).

[50] J. L. Cheng, J. E. Sipe, S. W. Wu, and C. Guo, APL Photon. **4**, 034201 (2019).