Beyond the effective mass approximation: A predictive theory of the nonlinear optical response of conduction electrons

Shukai Yu, Kate H. Heffernan, and Diyar Talbayev*

Department of Physics and Engineering Physics, Tulane University, 6400 Freret Street, New Orleans, Louisiana 70118, USA

(Received 5 July 2016; revised manuscript received 4 February 2017; published 9 March 2017)

We present an experimental and computational study of the nonlinear optical response of conduction electrons to intense terahertz (THz) electric field. Our observations (saturable absorption and an amplitude-dependent group refractive index) can be understood on the qualitative level as the breakdown of the effective mass approximation. However, a predictive theoretical description of the nonlinear THz propagation has been missing. We propose a model based on the semiclassical electron dynamics, a realistic band structure, and the free electron Drude parameters to accurately calculate the experimental observables in InSb. Our results open a path to modeling of the conduction-electron optical nonlinearity that governs the THz propagation in semiconductors.

DOI: 10.1103/PhysRevB.95.125201

I. INTRODUCTION

Studies of nonlinear optical response of conduction electrons in semiconductors have become possible in the last decade due to the advent of high-field terahertz (THz) sources [1–3] that opened a new frontier in nonlinear optics [4]. A rich variety of ultrafast nonlinear phenomena has been reported, such as THz-frequency modulation of semiconductor optical response [5–8], saturable absorption [9–12], self-phase modulation [13], ballistic motion of electrons [14,15], and interband tunneling and intervalley scattering of electrons [16–18]. Key features behind these observations are highly nonequilibrium excited states of electrons induced by the THz field and the breakdown of the effective mass approximation. While the emerging phenomenology of the nonlinear propagation is understood qualitatively, the quantitative connection between basic materials’ parameters and nonlinear THz propagation has been missing. A predictive theory of the optical nonlinearity poses important fundamental questions: while linear propagation is completely described by the Drude dielectric function, what parameters describe the nonlinear propagation? Which of the Drude parameters can be retained in the description of the nonlinear polarizability, and which ones must be abandoned? A quantitative description of nonlinear THz propagation is central to interpretation of the already discovered phenomenology. For example, increased intervalley scattering has been proposed in phenomenological interpretations that rely on effective linear optical constants to describe nonlinear propagation [9,18]. The predictive power of such phenomenological descriptions is limited. In addition, correct interpretations of the phenomenology must be based on a comparison of computed experimental observables that allows a separation of different contributions to nonlinearity, such as nonparabolicity of band structure, opening of new scattering channels, interband electronic excitation, etc. Interpretations based on effective linear optical constants do not achieve such separation.

In this article, we explore the optical nonlinearity due to conduction electrons in indium antimonide, InSb. Experimentally, we observe an amplitude-dependent group refractive index (a delay in the arrival time of the THz pulse) and saturable absorption (an increase in transmission at high incident field), which result from acceleration of conduction electrons to high crystal momenta and energies by THz electric field. We propose a model based on the realistic InSb band structure and semiclassical electron dynamics to account for the measured nonlinearity. The nonlinear polarizability is computed using the electron density $n$ and an energy-independent scattering rate $\gamma$ determined as Drude parameters from the linear optical properties. Keeping the scattering rate $\gamma$ independent of energy is not strictly justifiable when the electron samples large parts of the Brillouin zone. We make this assumption because of the simplicity that it affords. A computational implementation of the model using the finite-difference time-domain (FDTD) method provides a good quantitative agreement with experiment. Our main finding is the ability to use the linear Drude parameters, $n$ and $\gamma$, to accurately describe the nonlinear response. The effective mass approximation is replaced by the realistic band structure. Thus, we establish the connection between linear and nonlinear THz optical properties and provide a framework for predictive modeling of nonlinear THz propagation in other semiconductors [9,10,13,15,19–21].

II. EXPERIMENTAL DETAILS

The experiments were performed using a home-built THz spectrometer based on a 1 kHz repetition rate regenerative amplifier [22] and THz emission from a LiNbO$_3$ prism with tilted-wave-front phase matching [1–3]. We used electro-optic sampling [23] in ZnTe to estimate the peak THz electric field at the sample to be $\approx 100$ kV/cm. The sample was mounted on the cold finger of a closed-cycle He cryostat and was held at 10 K. The amplitude of the incident THz field on the sample was controlled by inserting Si attenuators in the THz beam path. In each measurement, a total of six Si attenuators were used and inserted in the parts of the spectrometer where the THz beam is collimated. To vary the THz field at the sample position, the attenuators were moved from before to after the sample in THz beam path. Each attenuator reduced the peak incident THz field by a factor of 0.56, with an almost flat frequency response. InSb is a low-band-gap semiconductor with a direct

*dtalbayev@gmail.com
band gap of 0.24 eV at 0 K [24], a low electron effective mass [25] \( m^* = 0.014 m_0 \), and a large nonparabolicity of the conduction band [26,27]. We used a slightly \( n \)-doped (nominally undoped) 0.5 mm thick (100) oriented InSb wafer for these measurements. The THz electric field was polarized along the (100) direction in the InSb crystal.

### III. RESULTS AND DISCUSSION

Figure 1(a) shows the measured evolution of the transmitted THz pulse with increasing peak amplitude of the incident THz field. The data are normalized to the peak incident electric field. Two nonlinear phenomena are apparent in the figure. First, the THz pulse arrives later at higher incident fields, the amplitude-dependent group delay. Second, the transmitted peak amplitude increases at higher incident fields, the saturable absorption. Nonlinear propagation of the THz pulse in InSb was studied in a \( z \)-scan measurement by Wen et al., who found a drop in transmission at the highest incident field [28]. They explain their findings by electron multiplication via impact ionization as the THz field accelerates electrons to energies above the band gap. The impact ionization was also observed in subsequent THz pump-probe studies of InSb [29]. Direct impact ionization is a very fast process that occurs on an \( \approx 140 \) fs time scale [30,31] with a very low threshold of the incident THz electric field [32] estimated to be \( \approx 8 \) kV/cm. Why is it not observed in our measurement? We conjecture that the amount of impact ionization may strongly depend on the initial density of electrons. Electron multiplication is a cascading process [28], as the accelerated electrons create new electrons that are also accelerated and create yet more electrons. In our sample, the measured electron density is \( n = 7.3 \times 10^{13} \) cm\(^{-3} \) at 10 K, which is considerably lower than the density cited in the previous studies [28,29]. At sufficiently low initial density, the impact ionization may have no appreciable effect on the propagation of the THz pulse, as evidenced by our experimental data.

Our data also show little to no evidence of interband electron tunneling due to the intense THz electric field. Interband electron tunneling was observed under high-field THz excitation in GaAs [16]. Electron tunneling would lead to an increase in conduction electron density and a drop in high-field THz transmission due to increased THz absorption. We only observe the transmission increase at the highest THz field, which is completely explained by the nonlinear electron dynamics at fixed electron density, as we will show below. The negligible electron tunneling rate in our experiments results from a combination of factors. First, the peak THz field in our measurement is 100 kV/cm, compared to 300 kV/cm reported for the observation of electron tunneling in GaAs [16], which partially offsets the possible increase in the tunneling probability due to the much lower band gap in InSb, 0.24 eV, compared to about 1.5 eV in GaAs. Second, the tunneling rate is proportional to the joint density of states at the edges of the conduction and valence bands, which in turn scales as \( (m^*)^{3/2} \). Due to the much lower electron and hole effective masses in InSb, the joint density of states is more than an order of magnitude lower in InSb than in GaAs. Finally, the electron tunneling rate due to THz field was reported by Kuehn et al. [16] to be proportional to the electron decoherence rate that scales linearly with momentum relaxation rate \( \gamma \), electron effective mass, and temperature. We estimate the electron decoherence rate in our measurement to be two orders of magnitude lower than in GaAs when the electron tunneling was observed [16].

The observed THz nonlinearity results from a breakdown of the effective mass approximation as electrons accelerate to highly nonparabolic parts of the conduction band. We model the nonlinear propagation computationally using the one-dimensional FDTD method [34], which relies on the Yee algorithm [35] to solve for the propagating THz fields \( E, D, \) and \( H \) inside InSb. For the linear propagation, the connection between the fields \( E \) and \( D \) is provided by

\[
D(\omega) = \varepsilon_0 \varepsilon(\omega) E(\omega)
\]

with the Drude dielectric function [36] \( \varepsilon(\omega) = \varepsilon_\infty (1 - \omega_p^2/\omega^2 + i\omega\gamma) \), where \( \varepsilon_\infty = 15.6 \) is the background...
dielectric constant and the plasma frequency $\omega_p^2 = \frac{n e^2}{\epsilon_0 \epsilon_\infty m^*}$ is related to the electron density $n$ and electron effective mass $m^*$. Equation (1) relies on the effective mass approximation and is no longer useful when the approximation breaks down. Instead, we compute the electric displacement as $\vec{D} = \epsilon_0 \epsilon_r \vec{E} + \vec{P}$, and the linear part of the displacement becomes $D^l = \epsilon_0 \vec{E} + \epsilon_0 \chi_0 \vec{E} = \epsilon_0 \epsilon_\infty \vec{E}$. We then add the nonlinear part $P^{NL}$ to this equation and obtain the relationship

$$D(r) = \epsilon_0 \epsilon_\infty \vec{E}(r) + P^{NL}(r).$$

We model the conduction electron response using the semi-classical description [37], in which the velocity of an electron wave packet of wave vector $\vec{k}$ is given by

$$v(k) = \frac{1}{\hbar} \frac{\partial \mathcal{E}(k)}{\partial k},$$

where $\mathcal{E}(k)$ is the conduction band energy-momentum dependence. The electron response to the electric field is governed by the equation of motion

$$\hbar \dot{k} + \gamma \hbar k = -e E(r, t),$$

where $\gamma$ is the electron scattering rate. Polarization $P^{NL}$ is computed as $P^{NL} = -n e v$, where $n$ is the electron density and the velocity $v$ must be determined from (3) using a realistic band structure of InSb after solving the equation of motion (4). Figure 2 shows the conduction band energy-momentum dependence $\mathcal{E}(k)$ and the velocity $v(k)$ in InSb used in our computation [33,38,39].

We combine Eqs. (2)–(4) to obtain a sequence of finite-difference equations that connect the $E$ and $D$ fields for the Yee algorithm. The electric field $E$ propagates along the $x$ direction and is polarized along $z$. The fields $E$ and $D$ only have $z$ components, as do the polarization $P^{NL}$ and the electron wave vector $k$. The field $H$ has only $y$ components. The FDTD algorithm computes the fields $H$, $D$, and $E$ for the time step $l + 1$ by using the field values from the earlier time steps, which are assumed known and stored in computer memory. We denote the time increment at each step as $\Delta t$ and write the time point for the $l$th step as $t_l = l \Delta t$. We first compute the fields $H$ and $D$ using Maxwell’s curl equations and the Yee central differencing in time and space for a nonpermeable medium [34]. This step does not yet include the connection to the conduction electron polarizability, which we establish next. From the equation of motion (4) we obtain an equation for the wave vector $k$ for the time step $l + 1$ at each point on the spatial grid:

$$k_{l+1} = (1 - \gamma \Delta t) k_l - (e \Delta t / \hbar) E_l.$$

After computing the wave vectors $k_{l+1}$, we determine the velocities $v_{l+1}$ from the realistic band structure (Fig. 2) and use them to compute the polarization and the electric field as

$$P^{NL}_{l+1} = P^{NL}_{l} - ne \Delta t (v_{l+1} + v_{l-1}),$$

$$E_{l+1} = (D_{l+1} - P^{NL}_{l+1}) / (\epsilon_0 \epsilon_\infty).$$

The Yee algorithm relies on finite-difference equations that are central about the time point $t_l$, and Eq. (5) is not. This is because the electric field value $E_{l+1}$ is not available to compute $k_{l+1}$ at this stage in the algorithm. To preserve the central difference nature of the algorithm, we compute $k_{l+1}$ again after the field $E_{l+1}$ is known using

$$k_{l+1} = \frac{1}{1 + \gamma \Delta t} k_{l-1} - \frac{e \Delta t}{\hbar (1 + \gamma \Delta t)} (E_{l+1} + E_{l-1}).$$

We then repeat steps (6) and (7). This completes the computation of the fields $E$, $D$, and $H$ for the time point $t_{l+1}$.

We used the above computational model to simulate the nonlinear propagation of the THz pulse in InSb. We determined the electron density $n$ and the scattering rate $\gamma$ from the linear THz spectroscopic measurement with very low incident electric field: $n = 7.3 \times 10^{13}$ cm$^{-3}$ and $\gamma = 0.5$ THz. The time increment was set to $\Delta t = 2.08$ fs and the space increment $\Delta x$ was set to 5 $\mu$m in vacuum and 1.25 $\mu$m inside InSb. We used two different incident source pulses: one was the realistic THz pulse recorded as the free-space reference in our measurement; the other was a Gaussian pulse given by $E(t) = E_0 \exp[-t^2/t_0^2] \sin[2\pi f_0 t]$ with $t_0 = 0.5$ ps and $f_0 = 1$ THz. The results of the computed nonlinear propagation of these THz pulses through a 0.5 mm InSb layer are shown in Figs. 1(b) and 1(c). Comparison with the experimental data [Fig. 1(a)] shows that the computation reproduces very well the major features of the nonlinear THz propagation—the group delay, the increase in the peak transmitted THz field, and the overall evolution of the pulse shape as the incident field gets stronger.

For a more quantitative comparison between the experiment and computation, we measured the change in the arrival time of the THz pulse (the group delay) with increasing peak THz field. We define the group delay as $\Delta \tau = (c \Delta \phi) / (2 \pi d)$, where $d$ is the sample thickness, $c$ is the speed of light, and $\Delta \phi$ is the average phase difference between high- and low-amplitude pulses transmitted by the sample in the frequency domain:

$$\Delta \phi = \arg \left[ \frac{E_{\text{high}}(\omega)}{E_{\text{low}}(\omega)} \right].$$

Here, we average over the full frequency content of the THz pulse. We also quantified the frequency-integrated
FIG. 3. Measured frequency-integrated transmission (circles) and group delay (diamonds) compared to the results of the simulated nonlinear THz propagation of the realistic (red line) and Gaussian (blue line) source pulses.

transmission as \[ T = \frac{\int E_{\text{sam}}^2(t)\,dt}{\int E_{\text{ref}}^2(t)\,dt}, \] (10)

where \( E_{\text{sam}} \) and \( E_{\text{ref}} \) are the time-domain electric fields transmitted by the InSb sample and the free-space reference. Figure 3 shows the experimental and computational group delay and integrated transmission. Our computational model describes very well the overall behavior of both parameters and provides a good quantitative agreement at the highest experimentally available peak THz fields. We emphasize that there are no free (fitting) parameters in the computational results of Fig. 3, as the electron density \( n \) and the scattering rate \( \gamma \) were fixed to the values obtained from the linear spectroscopic measurement. The only variable is the strength of the incident THz electric field.

We compared the experimental nonlinear propagation with THz electric field polarized along the \( \hat{z} \) and \( \hat{y} + \hat{z} \) directions and found no measurable difference. This observation is explained by a small maximum displacement of electrons in \( k \) space. Under the highest incident THz field, the maximum displacement reaches \( 0.1(2\pi/a) \). The conduction band is highly nonparabolic at these \( k \) values (Fig. 2), but the nonparabolicity is still isotropic [33], which is why we did not find any anisotropy of nonlinear THz propagation in the experiment.

We now look at the amplitude transmission \( T(\omega) = |\tilde{E}_{\text{sam}}(\omega)|/|\tilde{E}_{\text{ref}}(\omega)| \) in the frequency domain as a function of peak incident field (Fig. 4). Here, \( \tilde{E}_{\text{sam}}(\omega) \) and \( \tilde{E}_{\text{ref}}(\omega) \) are Fourier transforms of the time-domain wave forms transmitted by the sample and the reference aperture. The transmission measured in the linear regime at the lowest incident field (the blue line in all three panels in Fig. 4) is reproduced very well by the model. The linear transmission displays the characteristic dip below \( \sim 0.5 \) THz that corresponds to the Drude response of free electrons. In the absence of nonlinearity, all transmission curves in Fig. 4 would collapse onto the blue line. At high incident field in the nonlinear regime, we find deviations from the linear amplitude transmission. We also find differences between the experiment and the computational results; we find differences even between the realistic source and the Gaussian source computational results [Figs. 4(b) and 4(c)]. Despite the differences, the model captures well the main features of the nonlinear response, such as the increased transmission below 0.5 THz due to the saturation of the conduction electron absorption. The model also reproduces the nonmonotonic frequency dependence of the nonlinear transmission, which is apparent in Figs. 4(a) and 4(b): the transmission increases below 0.5 THz, decreases around 0.7 THz, and increases again above 1 THz. Another detail found in both the experiment and model is the nonmonotonic dependence of the nonlinear transmission on the incident THz amplitude at a specific frequency: for example, at 1.6 THz, the transmission first goes up and then goes down as the incident amplitude increases.

The difference between the nonlinear transmission computed using the realistic and Gaussian source pulses [Figs. 4(b) and 4(c)] does not come as a surprise. The realistic and Gaussian source pulses differ in their time domain shape and frequency content. The result of their nonlinear propagation cannot be obtained by applying the same transfer function...
to both pulses, as each Fourier frequency component will be enhanced or suppressed differently, depending on its amplitude and on the amplitudes of other Fourier components. Therefore, the different transmission for different time-domain inputs is a hallmark of nonlinear propagation. The same reasoning contributes to the difference in the measured and computed nonlinear transmission in Figs. 4(a) and 4(b). Even though the realistic source pulse in the model is the measured reference THz pulse, its measurement includes the response function of the THz receiver. The THz pulse that interacts nonlinearly with the InSb sample does not include the receiver response function and could have a slightly different time domain shape from the measured reference pulse. This difference causes no adverse effects in a linear sample-reference spectroscopic measurement, as the receiver response function cancels out when the linear transmission is computed. In our nonlinear case, the slight uncertainty in the shape of the incident pulse could contribute to the discrepancy in amplitude transmission between the measurement and the realistic source model [Figs. 4(a) and 4(b)].

The above pulse-shape uncertainty should not be viewed as the main or only source of the small discrepancies between theory and experiment. While mostly successful, our theory does not include such processes as interband electron tunneling or direct impact ionization, which could become important in different doping regimes or at higher incident THz fields. We also use an energy-independent scattering rate $\gamma$. As we argued earlier, interband tunneling and impact ionization are largely absent from our data, but could be present in incipient form and contribute to the observed nonlinear THz propagation. The same caveat applies to our assumption of the energy-independent scattering rate $\gamma$. The presented computational model is sufficiently flexible to allow the inclusion of additional physical processes via Monte Carlo–type ensemble descriptions of electron dynamics.

IV. CONCLUSIONS

The significance of our results is that a simple semiclassical model of electron dynamics, with the realistic band structure and the linear Drude parameters $n$ and $\gamma$, provides a good quantitative description of nonlinear THz propagation without the need to introduce phenomenological effective values of those parameters. We achieve good agreement between the computed and measured nonlinear absorption and refractive index (Fig. 3). The saturable absorption and the increased group refractive index have also been reported in other semiconductors [10,13], where they exhibit a very similar dependence on the strength of the incident THz field. This suggests that the validity of the proposed model should extend beyond the presented phenomenology of InSb. It is easily applied to other semiconductors by using the appropriate band structure. While we use here the average electron $k$ vector for simplicity, our method allows straightforward extensions to Monte Carlo–type descriptions of electron dynamics that make use of a distribution of electronic states [32]. As well, the method is sufficiently flexible to allow more sophisticated descriptions of electron scattering, such as those that introduce resonant intervalley scattering [9,10,31]. We proposed a computational method to establish a connection between microscopic electronic processes and the experimentally measured nonlinear THz propagation in semiconductors.

ACKNOWLEDGMENTS

This work was supported by the Louisiana Board of Regents Contracts No. LEQSF(2012-15)-RD-A-23 and No. LEQSF-EPS(2014)-PFUND-378, and by the NSF Award No. DMR-1554866.