High-frequency impact ionization and nonlinearity of photocurrent induced by intense terahertz radiation in HgTe-based quantum well structures

S. Hubmann,[1] S. Gebert,[1] G. V. Budkin,[2] V. V. Bel'kov,[2] E. L. Ivchenko,[2] A. P. Dmitriev,[2] S. Baumann,[1] M. Otteneder,[1] J. Ziegler,[1] D. Disterheft,[1] D. A. Kozlov,[3] N. N. Mikhailov,[3] S. A. Dvoretsky,[3] Z. D. Kvon,[3] D. Weiss,[1] and S. D. Ganichev[1]

[1]*Terahertz Center, University of Regensburg, 93040 Regensburg, Germany*
[2]*Ioffe Institute, 194021 St. Petersburg, Russia*
[3]*Rzhanov Institute of Semiconductor Physics, 630090 Novosibirsk, Russia*

We report on a strong nonlinear behavior of the photogalvanics and photoconductivity under excitation of HgTe quantum wells (QWs) by intense terahertz (THz) radiation. The increasing radiation intensity causes an inversion of the sign of the photocurrent and transition to its superlinear dependence on the intensity. The photoconductivity also shows a superlinear raise with the intensity. We show that the observed photoresponse nonlinearities are caused by the band-to-band *light* impact ionization under conditions of a photon energy less than the forbidden gap. The signature of this kind of impact ionization is that the angular radiation frequency $\omega = 2\pi f$ is much higher than the reciprocal momentum relaxation time. Thus the impact ionization takes place solely because of collisions in the presence of a high-frequency electric field. The effect has been measured on narrow HgTe/CdTe QWs of 5.7 nm width; the nonlinearity is detected for linearly and circularly polarized THz radiation with different frequencies ranging from $f = 0.6$ to 1.07 THz and intensities up to hundreds of kW/cm^2. We demonstrate that the probability of the impact ionization is proportional to the exponential function, $\exp(-E_0^2/E^2)$, of the radiation electric field amplitude E and the characteristic field parameter E_0. The effect is observable in a wide temperature range from 4.2 to 90 K, with the characteristic field increasing with rising temperature.

I. INTRODUCTION

With the emergence of high-power pulsed terahertz (THz) sources such as molecular lasers, free-electron lasers, and few-cycle difference-frequency based terahertz systems there has been a surge in studies of intense terahertz excitations of semiconductors, see, e.g., Refs. [1–9]. High-intensity electromagnetic radiation of the terahertz range gives rise to a variety of novel nonlinear phenomena whose characteristic features are basically different from the corresponding effects at microwave frequencies as well as for the visible light. In particular, high electric or magnetic fields of the THz radiation allow one a direct access to a number of low-energy elementary excitations such as phonons, plasma oscillations, spin waves, etc. or can drive the system into a nonperturbative regime of light-matter interaction. Moreover, the latest advances in terahertz technology made it possible to study nonlinear optical and optoelectronic phenomena on the femtosecond timescale with subcycle time resolution. Broadly speaking, experiments with powerful THz laser sources have potential to define limits of existing high-frequency electronics, where the radiation field has a classical amplitude, and uncover new approaches in the development and application of future electronics at these frequencies.

The intense THz radiation can be used to convert topologically trivial HgTe QWs into a nontrivial 2D Floquet topological insulator as suggested theoretically by Lindner *et al.* [10]. This would result in an appearance of chiral edge states which can be proved via the generation of edge photocurrents induced by circularly polarized THz radiation. The latter has recently been demonstrated in Ref. [11] studying 2D topological insulators based on HgTe QW of 8 nm width with inverted band ordering [12,13]. Being inspired by Ref. [10], we have investigated photoresponses in 5.7-nm-thick QWs excited by monochromatic intense THz radiation under the conditions suggested in that work. By examining the intensity dependence of the photogalvanic current, we have observed a strong nonlinearity resulting in a current sign inversion with increasing the radiation intensity I from a fraction of W/cm^2 up to hundreds of kW/cm^2 and a superlinear behavior at high power.

Further investigation and analysis have demonstrated, however, that the observed effect is caused by the light impact ionization [14]. This phenomenon is shown to cause the generation of electron-hole pairs by radiation with photon energy considerably smaller than the forbidden gap under the condition where the radiation angular frequency ω exceeds the reciprocal momentum relaxation time τ^{-1}. Under this condition, the charge carriers acquire high energies solely because of collisions in the presence of a high-frequency electric field [1,14,15]. Being primarily observed in bulk InSb crystals this effect was further demonstrated for very different three- and two-dimensional semiconductor systems [16–23]. A distinction of the light impact ionization reported in the present work is that it occurs in a system with the Fermi level larger than the forbidden gap. Thus, in contrast to all the previous works, the electron gas heating is needed in order to deplete the occupied states in the region of the

conduction-band bottom rather than to increase the number of hot electrons with high energies exceeding the energy gap. As shown below, the experiments and theoretical analysis demonstrate that in our study the probability of impact ionization is proportional to $\exp(-E_0^2/E^2)$, where E is the radiation electric field amplitude and E_0 is the characteristic field parameter. This dependence has been also confirmed by experiments on the THz radiation-induced photoconductivity showing that, in line with the theory, the characteristic field E_0 is proportional to the radiation frequency ω. Lastly, we show that the observed nonlinearity in the photocurrent is caused by the interplay of the photogalvanic effect excited in the conduction and valence bands.

II. SAMPLES AND METHODS

The nonlinear phenomena described here were studied in HgTe/HgCdTe QWs grown on (013)-oriented GaAs substrates by molecular beam epitaxy [24]. We used quantum wells with two barriers made of $Hg_{0.4}Cd_{0.6}Te$, each 30 nm thick. The width of the used QWs was $d = 5.7$ nm yielding a noninverted parabolic energy spectrum [12]. The structures have been grown on 4-μm-thick CdTe layers, which completely relax an initial strain caused by the lattice mismatch with the GaAs substrate. We used van-der-Pauw sample geometry with a size of 5×5 mm^2. To measure the photoresponse, six Ohmic contacts, four at the corners and two at the middle of opposite edges, have been fabricated. Besides the van-der-Pauw samples we fabricated Hall bar structures with a semitransparent gate. These structures were used to study dependence of the nonlinear photoresponse on the Fermi energy level position.

The 200-μm-wide and 450-μm-long Hall bars were patterned with photolithography and a Br-based wet etch process [25]. They are covered by an insulating oxide and a Ti/Au top gate electrode. The oxide layer consists of 30-nm SiO_2 and 100-nm Al_2O_3.

Additionally, Hall bar structures with a size of 6×50 μm have been prepared and used for magnetotransport measurements. From the latter measurements we obtained mobility μ and carrier density n at $T = 4.2$ K being $\mu = 2.5 \times 10^4$ cm^2 V^{-1} s^{-1} and $n = 3 \times 10^{11}$ cm^{-2}, respectively.

For optical excitation we used a high-power pulsed molecular gas THz laser [26–28] optically pumped by a tunable CO$_2$ laser [29]. Using CH$_3$F, D$_2$O, and NH$_3$ as active media, laser radiation with frequencies of 0.6, 0.77, and 1.07 THz ($\hbar\omega = 2.5$, 3.2, and 4.4 meV) were obtained. The laser operated in single pulse regime with a pulse duration of about 100 ns and a repetition frequency of 1 Hz. The radiation power P has been measured by a fast room temperature photon drag detector made of n-type Ge crystals [30]. The laser beam had an almost Gaussian shape as measured by a pyroelectric camera [31]. The radiation was focused by a parabolic mirror to a spot diameter of about 2.5 mm being smaller than the sample size. This allowed us to distinctly illuminate the edges or the center of the sample. In the following, the latter case is referred to as "bulk." The highest peak intensities obtained for these frequencies were 60, 80, and 200 kW/cm^2, respectively. Note that in the used laser the intensity varies from pulse to pulse by about 15%. The samples were placed in either an optical cold finger cryostat with TPX windows or an optical
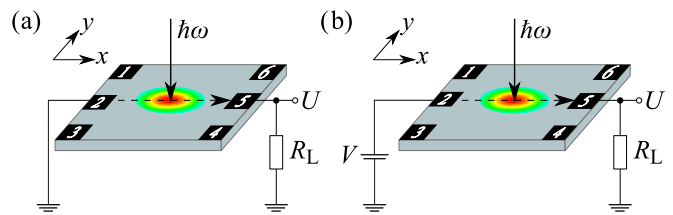

FIG. 1. (a) Setup scheme for a photocurrent measurement between contacts 2 and 5. The current is measured as a voltage drop U across a load resistor. (b) Setup scheme for a photoconductivity measurement between contacts 2 and 5. The signal is measured as a voltage drop U along a load resistor R_L, while a bias voltage V is applied.

temperature-regulated continuous flow cryostat with quartz windows. While in the former case we were able to obtain the highest level of radiation intensity at highest frequencies used in the work, in the latter case we studied the temperature dependence of the nonlinear response at the lowest frequency. The measurement have been carried out in the temperature range from $T = 4.2$ to 90 K. In all measurements the samples were illuminated at normal incidence, see Fig. 1.

To vary the laser radiation intensity, we used either calibrated attenuators or crossed polarizers. In the latter case, the linearly polarized laser radiation first passed through the wire grating polarizer. Rotation of this polarizer resulted in the decrease of the radiation intensity and the rotation of polarization state. The second polarizer, being at a fixed position, causes a further decrease of the radiation intensity and ensured that the radiation is always equally polarized. By this method, we obtained a controllable variation of the radiation intensity. To modify the radiation polarization state, crystal quartz quarter-wave plates were used. Rotating the plate by the angle φ with respect to the laser polarization along the x axis, we changed the degree of circular polarization P_{circ} according to

$$P_{circ} = \sin 2\varphi, \qquad (1)$$

and two other Stokes parameters P_L and $\tilde{P}_L$, defined as the degrees of linear polarization in the axes x, y and the axes x', y' rotated by 45°, according to [32]

$$P_L = \frac{\cos(4\varphi) + 1}{2}, \quad \tilde{P}_L = \frac{\sin(4\varphi)}{2}. \qquad (2)$$

Photocurrent and photoconductivity studied in this work have been measured using the setups shown in Figs. 1(a) and 1(b), respectively. The photocurrent was measured in *unbiased* samples via a voltage drop U across a load resistor R_L amplified by a voltage amplifier and detected by a digital broad-band oscilloscope. The dc photoconductivity was measured applying a bias voltage $V = 0.3$ V, see Fig. 1(b). This voltage was applied either as a dc voltage or as a pulsed bias with a pulse length much longer than the laser pulse, so that the voltage bias is in a "quasi-dc" regime with respect to the laser pulse. By subtracting the signals detected for positive and negative polarities of the bias, the photoconductivity signal was extracted. From this signal, the relative change of the conductivity $\Delta\sigma/\sigma$ has been calculated.
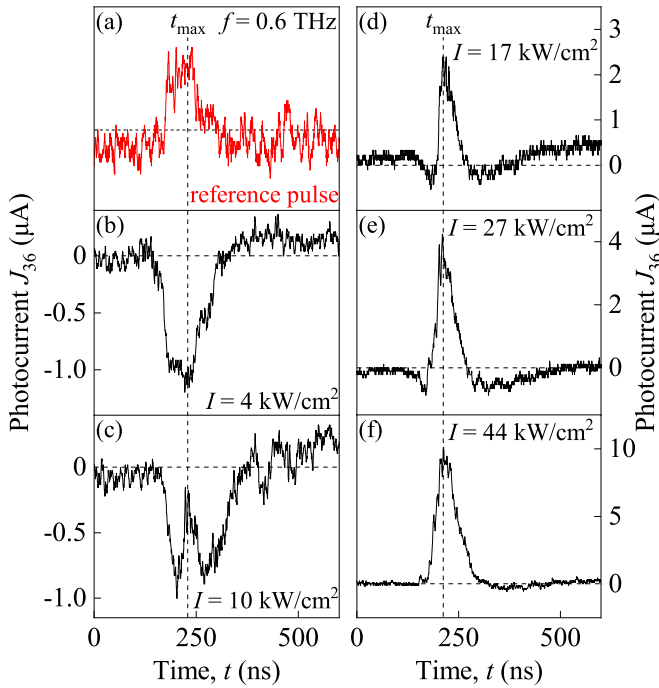
FIG. 2. (a) Reference laser pulse shape over time in arbitrary units. The signal is measured by the fast room temperature photon drag detector [30]. [(b)–(f)] Time variation of the photocurrent response excited by the illumination of the sample bulk with right-handed circularly polarized radiation at different intensities with a frequency of $f = 0.6$ THz and measured diagonally between contacts 3 and 6. At low intensities, the current response follows the reference pulse with a negative amplitude, while at the higher intensities a nonlinear positive contribution is observed, which results in a change of sign of the photocurrent at approximately 10 kW/cm^2.

III. RESULTS

A. Photocurrent

We firstly describe the results of the photocurrent measurements. Illuminating the edges or the bulk of the HgTe QWs with linearly or circularly polarized radiation we observed complex dynamics of the photoresponse at high radiation intensities. Figure 2 demonstrates an example of the genesis of the photoresponse with rising intensity detected for right-handed circularly polarized radiation with $f = 0.6$ THz. At relatively low incident intensities, the signal pulse is negative for any moment in time and its temporal shape repeats that of the reference laser pulse, Figs. 2(b) and 2(a), respectively. As the intensity is increased, the signal dynamic changes. Now, at first the negative signal amplitude increases, but then it rapidly drops to zero at the maximum of the excitation pulse ($t = t_{max}$), rises again and finally vanishes following the excitation pulse, see Fig. 2(c). On further increasing the intensity, the photosignal changes its sign at some value I_{inv} and becomes positive at the maximum of the excitation pulse. At even higher intensities, the positive part of the signal pulse dominates the photoresponse, Figs. 2(d)–2(f). The complicated temporal structure of the signal can be understood assuming that it emerges simply due to the change of sign of the photocurrent as a function of radiation intensity. On this assumption, at the rising edge of the laser pulse, the
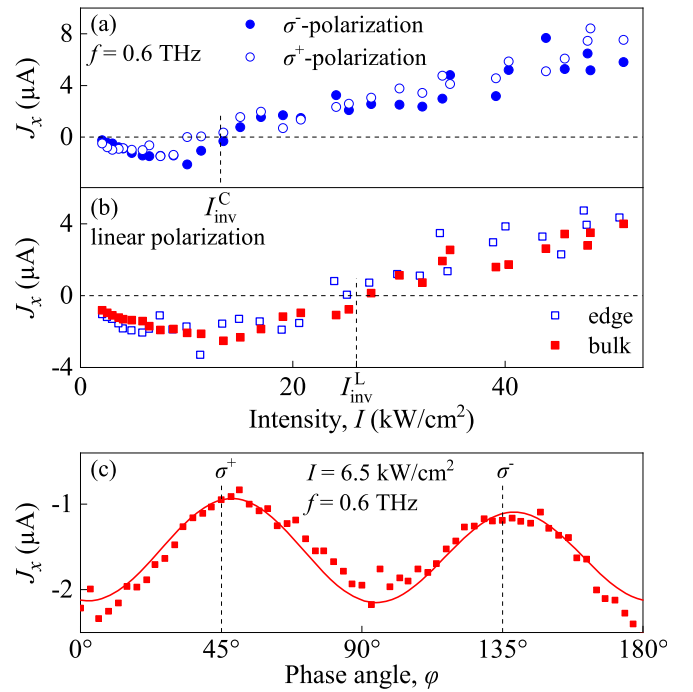

FIG. 3. (a) The intensity dependences of the edge photocurrent excited by right-handed (open circles) and left-handed (full circles) circularly polarized radiation with a frequency of $f = 0.6$ THz. The signal was picked up from contacts 1 and 6 under illumination of the corresponding sample edge. The change of sign and the nonlinear raise of the photocurrent is clearly seen, while there is no difference within the measuring accuracy between signals for the left- and right-handed excitation polarizations. (b) Dependencies of the photocurrent at the edge (blue, picked up from contacts 1 and 6) and in the sample bulk (red, picked up from contacts 2 and 5) excited by linearly polarized radiation with a frequency of $f = 0.6$ THz on the intensity I. (c) Dependence of the photocurrent induced in the sample bulk on the quarter-wave plate angle φ measured at the intensity of $I = 6.5$ kW/cm^2 for the radiation frequency of $f = 0.6$ THz. The solid line shows a fit after Eq. (3) with the fitting parameters $J_0 = -1.0 \, \mu$A, $J_1 = -1.1 \, \mu$A, $J_2 = -0.3 \, \mu$A, and $J_3 = 0.1 \, \mu$A.

intensity increase upon time causes the dynamic inversion of the photoresponse, whereas at the falling pulse edge, the intensity decreases and the signal dynamic mirrors. Consequently, the peak of the photocurrent occurs at the time t_{max} of the laser intensity maximum.

Figure 3 shows the intensity dependence of the photocurrent measured at time $t = t_{max}$. The open and full circles in Fig. 3(a) are obtained for illumination of the sample edge by circularly polarized radiation with opposite helicities. The data reveal almost no difference in the behavior of the photocurrents excited by the σ^+ and σ^- polarizations: the both dependences show the inversion of sign at the intensity of about $I_{inv}^C = 15$ kW/cm^2. The independence of the radiation helicity is additionally confirmed by studying the current variation with the phase angle φ (not shown). Similar intensity dependence, but with the higher value of $I_{inv}^L \approx 25$ kW/cm^2, has been observed for excitation with linearly polarized radiation, empty squares in Fig. 3(b). Studying the photoresponse in a wide temperature range from 4 to 90 K, we observed
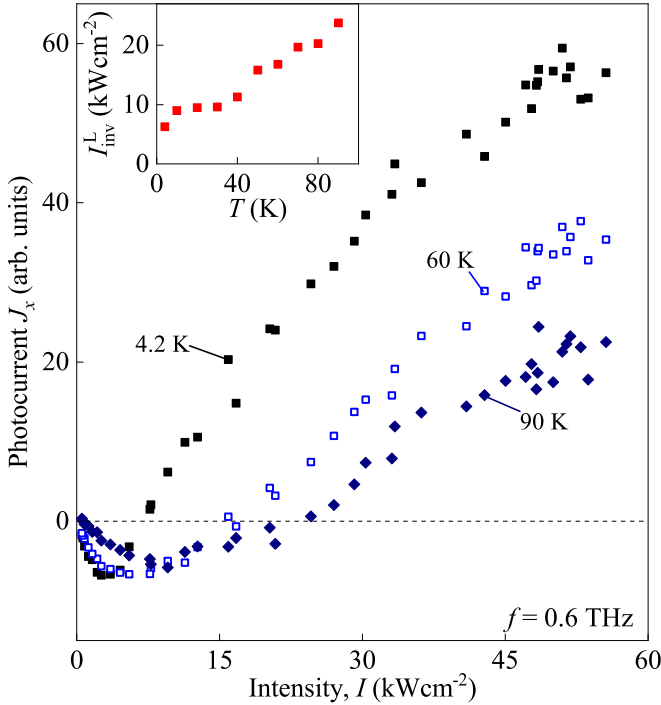
FIG. 4. Intensity dependencies of the photocurrent excited by linearly polarized radiation with a frequency of $f = 0.6$ THz for the temperatures of $T = 4.2$, 60, and 90 K. The signal was picked up from contacts 2 and 5 under illumination of the sample bulk. The inset shows the temperature dependence of the inversion intensity $I_{\text{inv}}^{\text{L}}$, which shifts to higher intensities at higher temperatures.

that qualitatively the nonlinear behavior remains the same, see Fig. 4. The only difference is that the intensity of the sign inversion increases with increasing temperature.

The insensitivity of the photoresponse to the radiation helicity demonstrates that the nonlinear photocurrent is not caused by helical edge photocurrents as a consequence of a transition from topologically trivial to nontrivial phase. Moreover, the current is not even generated at the sample edges: shifting the beam spot into the sample bulk (illumination of the sample center) and measuring the photocurrent across the sample we also observed the photocurrent sign change with rising the intensity. Figure 3(b) shows the comparison of the photocurrent excited by linearly polarized radiation at the edge (blue empty squares) and in the sample bulk (red full squares). The data reveal no measurable difference between the sample bulk and edge results excluding edge mechanisms as a driving force for the observed nonlinear photocurrent. In the following, we will present only the data obtained illuminating the bulk of the sample.

Furthermore, we have observed that switching of the radiation helicity from the right- to left-handed polarized radiation does not substantially affect the nonlinearity (not shown).

The photocurrent at low intensity is most likely caused by photogalvanic effects, well known for HgTe QWs [11,33]. This is supported by the polarization dependence detected at low intensities, see Fig. 3(c), which is described well by

$$J_x = J_0 + J_1(|e_x|^2 - |e_y|^2) + J_2\,\text{Re}(2e_x^*e_y) + J_3 P_{\text{circ}}$$
$$= J_0 + J_1 P_L(\varphi) + J_2 \tilde{P}_L(\varphi) + J_3 P_{\text{circ}}(\varphi), \quad (3)$$
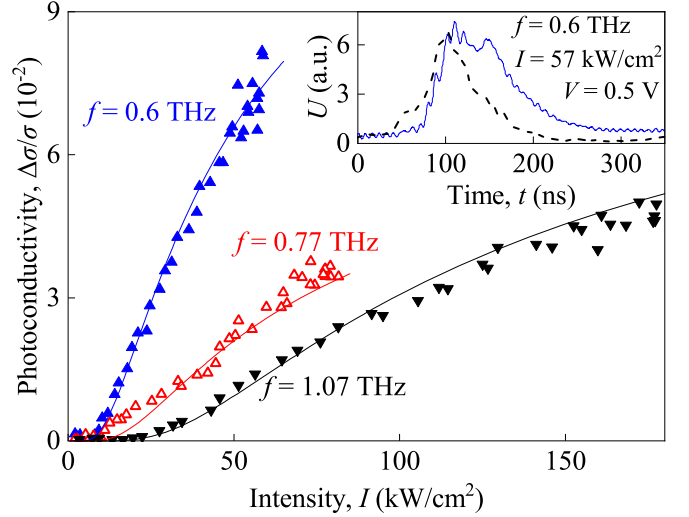

FIG. 5. Dependencies of the normalized photoconductivity $\Delta\sigma/\sigma$ excited by linearly polarized radiation in the bulk of the QW for the radiation frequencies of 0.6 (blue triangles), 0.77 (red triangles), and 1.07 THz (black triangles) on the intensity I. The signal is measured for contacts 2 and 5 using the setup shown in Fig. 1(b). Solid lines present the corresponding fits after Eq. (4) [see also theoretical Eq. (10)] with the fitting parameters A and I_0. A nonlinear raise of the signal is clearly seen for all frequencies. Inset shows the kinetics for an exemplary photoconductivity pulse (solid line) measured in the bulk of the samples for linearly polarized radiation with a frequency of 0.6 THz and an intensity of 57 kW/cm². The dashed line shows the reference laser pulse.

where e is the polarization unit vector and the Stokes parameters are defined by Eqs. (1) and (2). The fitting parameters J_0, J_1, and J_2 correspond to different contributions of the linear photogalvanic effect (LPGE) and the parameter J_3 describes the circular photogalvanic effect. Figure 3(c) reveals that the photocurrent is dominated by the LPGE.

As we show below (Sec. IV) the observed strong nonlinearity resulting in the change of the direction of the photogalvanic current with rising intensity is caused by the generation of electron-hole pairs by the THz radiation as a result of light impact ionization. To study the rate of electron-hole pair generation, we have measured the THz radiation induced change of conductivity $\Delta\sigma$.

B. Photoconductivity

A positive photoconductivity response was detected for all measured frequencies. In contrast to the photocurrent, the photoconductivity response time is twice longer than the laser pulse, see inset in Fig. 5. The observed increase of the sample's conductivity together with the photoresponse kinetics reveals that it is caused by the generation of electron-hole pairs, in spite of the fact that the photon energy is smaller than the band gap, $\hbar\omega < \varepsilon_{\text{g}}$. Note that the detected response times correspond to the recombination of electron-hole pairs known for HgTe QWs [34]. The characteristic intensity dependencies of the photoconductivity are shown in Fig. 5 for illumination of the sample bulk by linearly polarized radiation of different frequencies. Figure 6 shows the photoconductivity as a
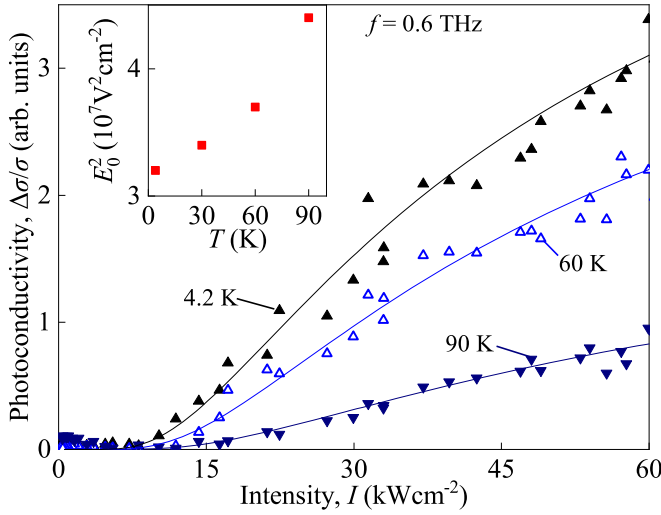
FIG. 6. Intensity dependences of the normalized photoconductivity $\Delta\sigma/\sigma$ excited by linearly polarized radiation with $f = 0.6$ THz in the bulk of the QW. The signal was measured for three temperatures of $T = 4.2$, 60, and 90 K for the contacts 2 and 5 using the setup shown in Fig. 1(b). Solid lines show corresponding fits after Eq. (4) [see also theoretical Eq. (10)] with the fitting parameters A and I_0. A nonlinear raise of the signal is clearly seen for all frequencies. Inset shows the temperature dependence of the fitting parameter $E_0^2 \propto I_0$.

function of the radiation intensity for the frequency of 0.6 THz and three temperatures in the range between 4 and 90 K.

The results clearly show a superlinear raise of the photoconductive signal and reveal that the nonlinearity decreases with increasing the radiation frequency. The data can be well described by

$$\frac{\Delta\sigma}{\sigma} = A \exp\left(-\frac{E_0^2}{E^2}\right) = A \cdot \exp\left(-\frac{I_0}{I}\right) \quad (4)$$

with the prefactor A and the characteristic electric field E_0 as fitting parameters. Replotting the data in a half-logarithmic plot as a function of the inverse squared electric field $E^{-2} \propto I^{-1}$ we see that the above equation describes the data well, see Fig. 7(a). Extracting E_0 from the slope of the fit lines we obtained that it changes according to $E_0 \propto f$, see inset in Fig. 7(b). We note that at low intensities and high frequencies a deviation from Eq. (4) occurs. This is attributed to a contribution of μ photoconductivity (bolometric photoresponse) [1]. Varying the Fermi energy by applying gate voltage in samples with a semitransparent gate we observed that reduction of ε_F increases the nonlinearity, see Fig. 8. The inset in Fig. 8 shows the dependence of the characteristic electric field E_0 on the Fermi level position. It demonstrates that E_0 almost linearly rises with the Fermi energy increase. Note that in the studied samples up to the highest possible negative gate voltages, the condition $\varepsilon_F \geqslant \varepsilon_g$ is satisfied.

To summarize the experimental part, we have shown that the intense terahertz excitation of HgTe QW structures results in a strongly nonlinear photoresponse which is caused by the generation of electron-hole pairs despite of the fact that the photon energy is substantially smaller than the energy gap. Electron-hole pair generation results (i) in a nonlinear
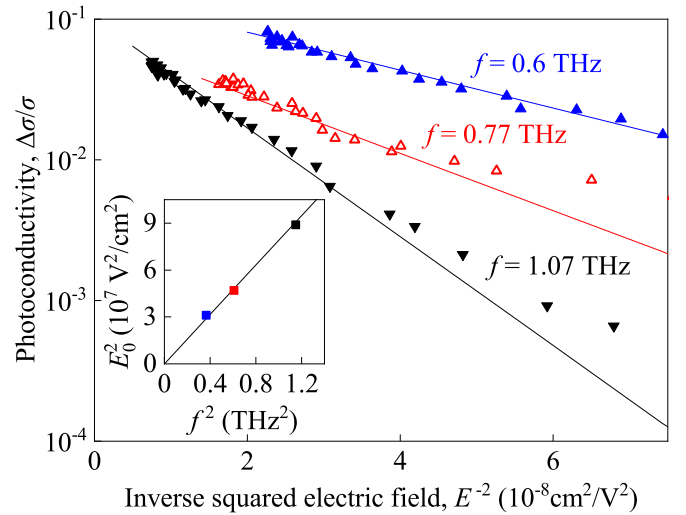

FIG. 7. Logarithmic plot of the relative photoconductivity $\Delta\sigma/\sigma$ excited by linearly polarized radiation in the bulk of the sample on the inverse squared electric field E^{-2}. Data are shown for the radiation frequencies of 0.6 (blue triangles), 0.77 (red triangles), and 1.07 THz (black triangles) and have been measured between the contacts 2 and 5. Solid lines present the corresponding fits after Eq. (4) [see also theoretical Eqs. (10) and (11).] with fitting parameters A and E_0^2. Inset shows a dependence of the fitting parameter E_0^2 extracted from the corresponding fits from panel (a) on the squared radiation frequency f^2.

photoconductivity, which scales as $\exp(-E_0^2/E^2)$ with $E_0 \propto f$, and (ii) in a sign inversion of the photogalvanic currents at high intensities. The fact that for smaller photon energies the nonlinearity is observed at substantially smaller radiation intensities, see Fig. 7, together with the lack of the linear-circular dichroism (not shown), excludes multiphoton processes as an origin of the nonlinearities. The observed electric field and frequency dependencies of the photoconductivity give a hint that the observed nonlinearity is caused by the light impact ionization, previously observed in bulk InSb and InAs-based materials [14,16,17,35].

IV. THEORY AND DISCUSSION

Now we discuss the microscopic origin of the observed nonlinear photoconductivity and photocurrent. We begin with the discussion of the band structure of the studied samples which is required to understand the mechanism of the electron-hole pair generation. The band structure of 3D HgTe is related to the work of Groves *et al.* [36]. Here we calculate electron states and dispersion of the (013)-oriented 5.7-nm-thick HgTe QW in the eight-band model including the conduction (Γ_6), valence (Γ_8), and spin-orbit split-off bands (Γ_7). The effective $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian in the second-order Kane model, see, e.g., Ref. [37], is taken from Ref. [38]. Figure 9 shows the energy dispersion for three subbands $E1$, $H1$, and $H2$. Here, at $k_\parallel = 0$, the $|E1, \pm 1/2\rangle$ quantum well subband state is formed from the linear combination of the $|\Gamma_6, \pm 1/2\rangle$ and $|\Gamma_8, \pm 1/2\rangle$ states, while $|H1\rangle$ and $|H2\rangle$ states are the first and second levels of the size quantization of heavy holes, respectively. Figure 9 reveals that the band gap in the studied QWs is
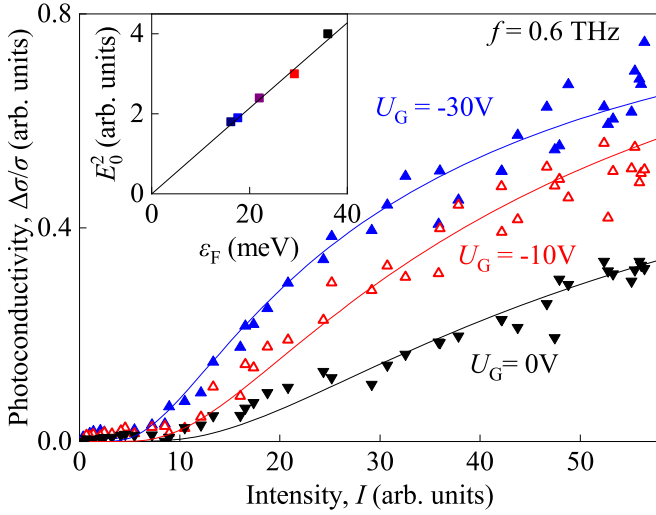
FIG. 8. Intensity dependencies of the normalized photoconductivity $\Delta\sigma/\sigma$ excited by linearly polarized radiation with $f = 0.6$ THz in the sample with semitransparent gate. Solid lines show corresponding fits after Eq. (4) [see also theoretical Eq. (10)] with the fitting parameters A and $I_0 \propto E_0^2$. Inset shows the dependence of the fitting parameter E_0^2 on the Fermi energy. The latter has been obtained from the corresponding magnetotransport measurements. Note that due to unknown intensity acting on 2DEG after transmission through the gate all values are given in arbitrary units.

$\varepsilon_g = 17.6$ meV. From magnetotransport measurements we obtained that the Fermi level is $\varepsilon_F = 54$ meV, i.e., is substantially larger than the band gap.

A. Photoconductivity

In all previous works aimed to impact ionization, the typical values of the electron energy ε in the equilibrium state are much lower than the threshold energy of impact ionization ε_i. Therefore, to achieve an appreciable rate of generation of the pairs, a sufficiently strong heating of the electron gas is
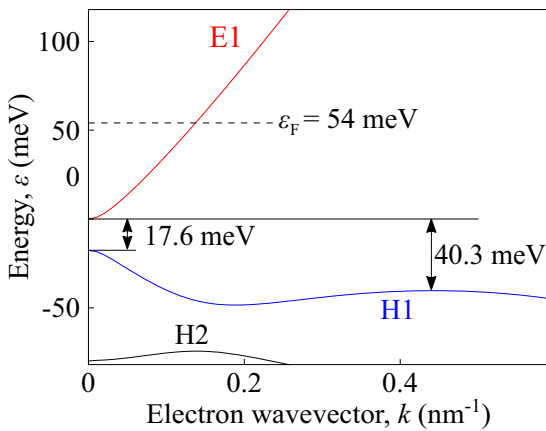

FIG. 9. Calculated energy dispersion for a quantum well with a thickness of 5.7 nm using the $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian from Ref. [38]. Here, the band gap between the electron subband E1 (red) and the hole subband H1 (blue) is $\varepsilon_g = 17.6$ meV. The dashed line indicates the Fermi level.

necessary, leading to an increase in the number of electrons with high energy.

In the samples studied here at rather high electron concentrations, the Fermi energy ε_F exceeds ε_i already in the equilibrium. In our case ε_i is about half as much as ε_F. Thus the generation of pairs is limited not by the lack of electrons with high energy, but rather by the small number of unoccupied levels in the low-energy region. In such a situation, heating is needed in order to deplete the occupied levels in the region of relatively low energies.

We consider the dominant mechanisms of electron scattering in the discussed polar HgTe crystals at low temperatures to be spontaneous emission of polar optical phonons with the energy $\varepsilon_0 = \hbar\omega_0$ and elastic scattering by point impurities with time $\tau_i(\varepsilon)$. We also take into account that for the experiments discussed above the condition $\omega\tau > 1$ is fulfilled. Indeed, from the width of the cyclotron resonance, we obtained $\omega\tau \approx 2$ for the studied range of frequencies.

For a radiation electric field polarized along the x axis relevant to the experiment, the kinetic equation has the form

$$\frac{\partial f(\boldsymbol{p}, t)}{\partial t} + eE\cos(\omega t)\frac{\partial f(\boldsymbol{p}, t)}{\partial p_x}$$
$$= -\frac{f(\boldsymbol{p}, t) - f_0(\varepsilon, t)}{\tau_i(\varepsilon)} - St_-^{ph}\{f\} + St_+^{ph}\{f\}. \quad (5)$$

Here, $f(\boldsymbol{p}, t)$ is the distribution function depending on electron momentum $\boldsymbol{p}$ and time t, e is the electron charge, $p(\varepsilon)$ and $v(\varepsilon) = \partial\varepsilon/\partial p$ are the momentum and velocity of an electron with the energy ε; $f_0(\varepsilon, t)$ is the part of the distribution function that is independent of the direction of the electron momentum $\boldsymbol{p}$; $\tau_i(\varepsilon) = \tau_{i0}g(0)/g(\varepsilon)$, where τ_{i0} is scattering time at the bottom of the conduction band [39,40]; $g(\varepsilon) = p(\varepsilon)/v(\varepsilon)\pi\hbar^2$ is the density of states in the conduction band.

The combination $St_+^{ph}\{f\} - St_-^{ph}\{f\}$ describes the collision integral due to interaction with optical phonons. Separately, $St_+^{ph}\{f\}$ ($St_-^{ph}\{f\}$) describes the number of electrons entering (leaving) the state with momentum $\boldsymbol{p}$ in an unit of time due to emission of phonons.

For the considered dominant mechanisms of electron scattering, the distribution function $f(\boldsymbol{p}, t)$ is almost isotropic, i.e., $f(\boldsymbol{p}, t) = f_0(\varepsilon) + f_1(\boldsymbol{p}, t)$, where $f_1(\boldsymbol{p}, t)$ is a small anisotropic correction. We assume that the electron gas heating is sufficiently strong, so that the effective electron temperature exceeds the phonon energy and, therefore, the function $f_0(\varepsilon)$ only slightly changes on the energy scale of the order of ε_0. By that for $\varepsilon > \varepsilon_0$, $f_0(\varepsilon)$ is given by the balance equation

$$D(\varepsilon)\frac{\partial f_0}{\partial \varepsilon} + \frac{\varepsilon_0}{\tau_{ph}(\varepsilon)}f_0(\varepsilon)[1 - f_0(\varepsilon)] = 0,$$
$$D(\varepsilon) = \frac{e^2 E^2 v^2(\varepsilon)}{4\omega^2 \tau_i(\varepsilon)}, \quad \frac{1}{\tau_{ph}(\varepsilon)} = \frac{4\pi\varepsilon_0 e^2 g(\varepsilon)}{\bar{\epsilon}p(\varepsilon)}, \quad (6)$$

where $1/\bar{\epsilon} = 1/\epsilon_\infty - 1/\epsilon_0$; ϵ_∞ and ϵ_0 are high- and low-frequency dielectric permittivities; $D(\varepsilon)$ is the diffusion coefficient of electrons in the energy space, and $\tau_{ph}(\varepsilon)$ is the characteristic time of phonon emission. The first term in Eq. (6) describes the heating of the electron gas by the electric field of the electromagnetic wave, whereas the second term corresponds to the energy losses due to emission of phonons.

The solution of Eq. (6) has the form

$$f_0(\varepsilon) = \frac{1}{1 + \exp[-L(\varepsilon)]}, \quad L(\varepsilon) = \int_\varepsilon^{\varepsilon_E} \frac{\varepsilon_0}{D(\varepsilon')\tau_{ph}(\varepsilon')} d\varepsilon',$$

(7)

where ε_E is determined from $n = \int_0^\infty f_0(\varepsilon)g(\varepsilon)d\varepsilon$, i.e., normalization by the density n.

For the considered condition of a slowly changing function $f_0(\varepsilon)$ on the energy scale of the order of ε_0, the value $1 - f_0(\varepsilon)$ also varies only slightly with energy ε. At $f_0(\varepsilon)$ close to unity, $1 - f_0(\varepsilon)$ is proportional to $\exp[-L(\varepsilon)]$. The above condition is fulfilled for $|L(\varepsilon + \varepsilon_0) - L(\varepsilon)| \ll 1$, which as it follows from Eq. (7) is equivalent to $\varepsilon_0^2 \ll D\tau_{ph}$. For the opposite inequality, $\varepsilon < \varepsilon_0$, the distribution function obeys the equation

$$-\frac{1}{g(\varepsilon)} \frac{\partial}{\partial \varepsilon}\left[g(\varepsilon)D(\varepsilon)\frac{\partial f_0}{\partial \varepsilon}\right] = \frac{1}{t_{ph}^+(\varepsilon)} f_0(\varepsilon + \varepsilon_0)[1 - f_0(\varepsilon)],$$

(8)

where $f_0(\varepsilon + \varepsilon_0)$ is given by Eq. (6) and t_{ph}^+ is given by Eq. (A12) in Appendix.

While it is not possible to solve this equation analytically, under the conditions $\varepsilon_0^2 \ll D\tau_{ph}$ and $f_0(\varepsilon) \approx 1$ relevant to the considerate case the right part of Eq. (8) becomes vanishingly small and, therefore, the distribution function in the region $\varepsilon < \varepsilon_0$ can be considered as almost constant. It can be calculated from Eq. (7) taking into account $\varepsilon = \varepsilon_0$. For a not so strong electron gas heating, so that the effective electron temperature is smaller than the Fermi energy, the distribution function is close to unity in the whole range of energies $0 < \varepsilon < \varepsilon_E$. By contrast, at high energies $\varepsilon \gg \varepsilon_E$, it approaches zero. Consequently one obtains that $n \approx \int_0^{\varepsilon_E} g(\varepsilon)d\varepsilon$, which for ε_E yields the value equal to the Fermi energy in equilibrium. This is because for the degenerated electron gas the latter is determined by the same integral.

The ionization rate is determined, as stated before, by the number of unoccupied levels in the low-energy region, i.e., by $1 - f_0(\varepsilon) \approx \exp[-L(\varepsilon)]$. The exact value of the lower limit of integration in the expression for $L(\varepsilon)$ depends on the characteristics of the elementary act of impact ionization unknown to us, and therefore we shall consider it to be an adjustable parameter ε^*. Subsequently, field and frequency dependencies of the number of generated pairs are determined by the exponent $\exp[-L^*]$, where

$$L^* = \int_{\varepsilon^*}^{\varepsilon_F} \frac{\varepsilon_0}{D(\varepsilon)t_{ph}(\varepsilon)} d\varepsilon.$$

(9)

Within the Bernevig-Hughes-Zhang model [12], the energy is described by $\varepsilon(p) = \sqrt{\varepsilon_g^2/4 + p^2\varepsilon_g/2m} - \varepsilon_g/2$, which yields $g(\varepsilon) = m(2\varepsilon + \varepsilon_g)/\pi\hbar^2\varepsilon_g$, $p(\varepsilon) = \sqrt{2m\varepsilon(\varepsilon + \varepsilon_g)\varepsilon_g^{-1}}$ and $v^2(\varepsilon) = 2\varepsilon(\varepsilon + \varepsilon_g)\varepsilon_g/m(2\varepsilon - \varepsilon_g)^2$. Here, m is the effective mass of the electron at the bottom of the conduction band. Substituting these expressions into Eq. (6), we obtain the rate of ionization:

$$W = W_i \exp(-L^*) = W_i \exp\left(-\frac{E_0^2}{E^2}\right)$$

(10)

with

$$E_0^2 = \frac{8\pi\omega_0^2\omega^2 m^{3/2}\tau_{i0}}{\bar{\epsilon}\sqrt{2\varepsilon_g}} \int_{\varepsilon^*/\varepsilon_g}^{\varepsilon_F/\varepsilon_g} \frac{(2z + 1)^2}{(z^2 + z)^{3/2}} dz,$$

(11)

where W_i is the probability of the single impact ionization event. This equation describes well the main features of the observed photoconductivity. Indeed, as discussed above, at high electric fields the relative change of the conductivity $\Delta\sigma/\sigma \propto W$ varies exponentially with the square of the inverse radiation electric field, see Fig. 7. Furthermore, the above equation reveals that the characteristic electric field E_0 scales linearly with the radiation angular frequency ω, which also corresponds to the results of the experiments, see inset in Fig. 7. The observed increase of the E_0^2 with the increase of the Fermi energy is also in a qualitatively agreement with Eq. (11). Note that a quantitative description of the Fermi energy dependence is not an easy task because with the increase of the upper limit of the integration the lower limit, being a fitting parameter in our calculations, will change as well. Finally we note that the observed increase of the characteristic field E_0 with rising temperature, see inset in Fig. 6, is not surprising, showing that the generation rate of electron-hole pairs decreases with increasing temperature. Indeed, typically at higher temperatures, the same relative increase of the electron temperature is obtained at higher electric fields.

B. Photocurrent

Now we discuss the observed photocurrent and its non-linearity at high intensities. Instead of the quasimomentum $\boldsymbol{p}$, we use the wave vector $\boldsymbol{k} = \boldsymbol{p}/\hbar$ for the description of the charge carrier states. We start with the analysis of the photocurrent formation mechanism. As addressed above it is mainly caused by the LPGE, in the experiment under study the circular photocurrent contribution is unessential. Two mechanisms can contribute to the LPGE current generation: (i) the shift photocurrent which arises from the second-order interaction with the electromagnetic wave and is related to the carrier displacement in the real space under direct or indirect optical transitions [33,41–47]; and (ii) the ballistic photocurrent, which arises beyond the Born approximation in calculating the optical matrix element. In the latter case, the carrier transition rate is asymmetric in the $\boldsymbol{k}$ space, and the photocurrent is stabilized by the momentum scattering time τ [42,44,48].

In general, both contributions appear due to the same asymmetries of elementary processes of photon absorption and charge carrier scattering by phonons or defects and are comparable in order of magnitude. Here we fix the attention on the shift contribution and employ a model which describes solely the LPGE.

We consider the photocurrent generated under intraband indirect (Drude-like) optical transitions in the lowest conduction band (c) and the highest valence band (v). Taking into account that in our experiment at high intensities both electrons and holes are simultaneously present in the sample, the shift photocurrent density is a sum $\boldsymbol{j}^{(c)} + \boldsymbol{j}^{(v)}$ of the electron and hole contributions. The shift photocurrent density $\boldsymbol{j}^{(l)}$ ($l = c, v$) is given by the following microscopic

equation [44]:

$$j^{(l)} = 2q_l \sum_{k'k} W_l(k', k) R_l(k', k), \qquad (12)$$

where $W_l(k', k)$ is the probability rate of the optical transitions from the free-carrier state $|lk\rangle$ to $|lk'\rangle$, q_l is the charge of free carriers, $q_c = e$ for an electron and $q_v = -e = |e|$ for a hole, $R_l(k', k)$ is the elementary shift under an indirect one-photon transition, and the factor of 2 accounts for the spin degeneracy.

The intraband photon absorption process involves participation of a static defect or a phonon. In the limit $\omega \gg \tau^{-1}$, the transition rate is calculated using Fermi's golden rule for indirect transitions. For static-defect-assisted transitions, this rule reads

$$W_l(k', k) = \frac{2\pi}{\hbar} N_d |M_{lk',lk}|^2 (f_{lk} - f_{lk'}) \delta(E_{lk'} - E_{lk} - \hbar\omega), \qquad (13)$$

with the compound matrix element being

$$M_{lk',lk} = \frac{V_l(k) - V_l(k')}{\hbar\omega} U_l(k', k). \qquad (14)$$

Here, $U_l(k', k)$ is the matrix element of scattering by a defect, m_l is the effective mass m_c or m_v, N_d is density of scattering defects, $E_{lk} = \hbar^2 k^2/(2m_l)$, f_{lk} is the electron distribution function, $V_l(k)$ is the matrix element of carrier-light interaction, the difference of the matrix elements can be presented as

$$V_l(k) - V_l(k') = \frac{q_l \hbar A_0}{m_l c} e \cdot (k' - k), \qquad (15)$$

where A_0 is the scalar amplitude of the light vector potential. While deriving Eq. (14) we took into account the energy conservation law $E_{lk'} - E_{lk} = \hbar\omega$. Using the theory of the shift LPGE [45] derived for the indirect optical transitions, we obtain for the elementary shift in Eq. (12)

$$R_l(k', k) = -\frac{\text{Im}[U_l^*(k', k)(\nabla_{k'} + \nabla_k) U_l(k', k)]}{|U_{lk',lk}|^2} + \mathcal{A}_{lk'} - \mathcal{A}_{lk}, \qquad (16)$$

where $\mathcal{A}_{lk}$ is the Berry connection. In the steady-state regime of photoexcitation, the contributions of the Berry connections from all the processes cancel each other. Because of the hermiticity of the operator of carrier-defect interaction, one has

$$U_l^*(k', k) = U_l(k, k'). \qquad (17)$$

It is worth to note that if the matrix element $U_l(k', k)$ depends only on the difference $k' - k$ the shift (16) vanishes. Therefore we have to make allowance for the dependence of $U_l(k', k)$ not only on the difference but also on the sum $k' + k$. For brevity, we shall leave aside the general case and discuss a special model leading to a nonvanishing shift (16). Namely, we expand the matrix element (17) up to the second order as follows:

$$U_l(k', k) = U_0^{(l)} + i U_{1,\alpha}^{(l)}(k'_\alpha - k_\alpha) + U_{2,\beta\gamma}^{(l)}(k'_\beta k'_\gamma + k_\beta k_\gamma), \qquad (18)$$

where $\alpha, \beta, \gamma = x, y$ and $U_{2,\beta\gamma}^{(l)} = U_{2,\gamma\beta}^{(l)}$. The shift contributing to the photocurrent is given by

$$|U_l(k', k)|^2 R_{l\beta}(k', k) = 2 U_{1,\alpha}^{(l)} U_{2,\beta\gamma}^{(l)}(k'_\alpha - k_\alpha)(k'_\gamma + k_\gamma). \qquad (19)$$

We omit intermediate calculations and give final formulas for the photocurrents. For the linearly polarized light ($E \parallel x$) passed through the $\lambda/4$ plate, the result reads

$$j_{l,x} = \left(\chi_{l,x}^{(0)} + \chi_{l,x}^{(1)} \frac{1 + \cos 4\varphi}{2} + \chi_{l,x}^{(2)} \frac{\sin 4\varphi}{2}\right) n_l I,$$

$$j_{l,y} = \left(\chi_{l,y}^{(0)} - \chi_{l,y}^{(1)} \frac{1 + \cos 4\varphi}{2} + \chi_{l,y}^{(2)} \frac{\sin 4\varphi}{2}\right) n_l I, \quad (20)$$

in agreement with Eq. (3). Here, n_l are the electron and holes densities, n for $l = c$ and p for $l = v$, coefficients $\chi_{i,\alpha}^{(l)}$ are given by

$$\chi_{l,\alpha}^{(0)} = 2 \frac{\eta_l(I)}{n_l} \frac{q_l m_l}{\hbar^2} \frac{U_{2,\alpha\gamma}^{(l)} U_{1,\gamma}^{(l)}}{U_0^{(l)2}},$$

$$\chi_{l,\alpha}^{(1)} = \frac{\eta_l(I)}{n_l} \frac{q_l m_l}{\hbar^2} \frac{U_{2,\alpha\gamma}^{(l)} U_{1,\gamma}^{(l)}(2\delta_{\alpha\gamma} - 1)}{U_0^{(l)2}}, \qquad (21)$$

$$\chi_{l,\alpha}^{(2)} = \frac{\eta_l(I)}{n_l} \frac{q_l m_l}{\hbar^2} \frac{U_{2,\alpha\beta}^{(l)} U_{1,\gamma}^{(l)}(1 - \delta_{\beta\gamma})}{U_0^{(l)2}},$$

where $\eta_l(I)$ is the absorbance defined by

$$\eta_l(I) = W_l(I) \frac{\hbar\omega}{I}, \qquad (22)$$

and $W_l(I)$ is the absorption rate per unit area contributed by the charge carriers l. Because the electron and hole densities are nonequilibrium the absorbance is dependent on the light intensity. While deriving Eqs. (21) we assumed the k-dependent terms in the expansion (18) to be smaller as compared with the first term $U_0^{(l)}$. The analysis shows that in the experimental conditions related to Fig. 10 the ratio $\eta_l(I)/n_l$ is independent of n_l and the introduction of concentration as a factor in Eqs. (20) is justified.

The (013)-oriented HgTe quantum-well structure has no point-group symmetry elements except for the trivial identity operation C_1. As a result, the structure has no special in-plane axes, the symmetry impose no restrictions on the coefficients $U_{1,\alpha}^{(l)}$ and $U_{2,\alpha}^{(l)}$ in Eq. (18) and that, in Eq. (20), the sets of coefficients $\chi_{c,\alpha}^{(i)}$ for the conduction electrons and coefficients $\chi_{v,\alpha}^{(i)}$ for the holes are linearly independent. This means, particularly, that the electron and hole photocurrents are not necessarily parallel to each other, see the final remarks at the end of the section. Equations (20) describe well the experimentally observed polarization dependences of the photocurrent $J \propto j$ as shown in Fig. 3(c). To fit experimental data, we consider polarized radiation with electric field vector E polarized along the x axis and the photosignal measured along the x direction, between the contacts 2 and 5.

Now we discuss the observed nonlinearity and dynamical sign change of the photocurrent. The process of light impact ionization creates electron-hole pairs resulting in an increase of the photogalvanic effect in the conduction band
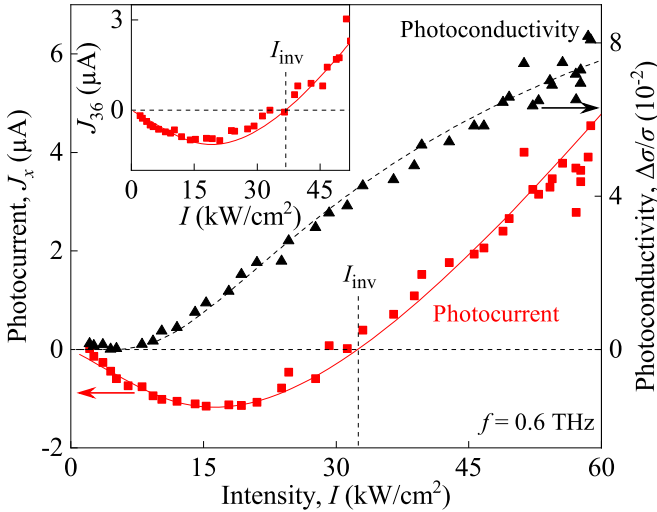
FIG. 10. Dependencies of the photoconductivity (black triangles) and the photocurrent (red squares) excited by linearly polarized radiation with a frequency of $f = 0.6$ THz on the intensity. Signals were measured between contacts 2 and 5. Dashed black line shows fit after Eq. (4). Here the fitting parameters $A = 0.15$ and $I_0 = 41$ kW/cm^2 obtained from the photoconductivity measurements, see Figs. 5 and 7, are used. Solid red line shows a fit according to Eq. (25) [also see Eqs. (23) and (24)]. Here, A and I_0 are taken from the photoconductivity fit, $n = 3 \times 10^{11}$ cm^{-2} from transport measurements and the parameters $\chi_c = 3.4 \times 10^{-13}$ μA cm^4 kW^{-1} and $\chi_v = 8.2 \times 10^{-12}$ μA cm^4 kW^{-1} are obtained from fitting. At low intensities, the photocurrent varies linearly with the intensity, while the photoconductivity signal is negligibly small. At higher intensities the photocurrent starts to show nonlinear behavior as the photoconductivity increases superlinearly. The inset shows the dependence of the diagonal photocurrent between contacts 3 and 6 excited by linearly polarized radiation with a frequency of $f = 0.6$ THz on the intensity. The solid line shows a fit according to Eq. (25) with the fitting parameters $\chi_c = 2.7 \times 10^{-13}$ μA cm^4 kW^{-1} and $\chi_v = 6 \times 10^{-12}$ μA cm^4 kW^{-1}. It is seen that the inversion intensity I_{inv} increases in this measurement direction.

described by

$$j_x^{(c)} = \chi_c(n_0 + \Delta n)I, \qquad (23)$$

as well as in the generation of holes and appearance of a hole photogalvanic current described by

$$j_x^{(v)} = \chi_v \Delta p I. \qquad (24)$$

Here, n_0 is the equilibrium electron density, $\Delta n = \Delta p \propto \exp(-I_0/I)$ are the densities of electrons and holes generated by the light impact ionization,

$$\chi_c = \kappa\left(\chi_{c,x}^{(0)} + \chi_{c,x}^{(1)}\right), \quad \chi_v = \kappa\left(\chi_{v,x}^{(0)} + \chi_{v,x}^{(1)}\right),$$

κ is the parameter that relates the measured photocurrent to the current density (20) and is determined by the sample geometry and the laser spot position and diameter. With increasing the intensity the hole density increases and, in case $|\chi_v| > |\chi_c|$, the valence band contribution of the photogalvanic effect may become comparable to and even exceed that of the conduction band.

Fitting our data by the sum

$$j_x = j_x^{(c)} + j_x^{(v)} \qquad (25)$$

and assuming the contributions of electrons and holes to have opposite signs, we have obtained a good agreement taking the value of χ_v by an order of magnitude larger than χ_c, see Fig. 10. In this figure, the data are fitted using only one adjustable parameter χ_v/χ_c. The coefficient A determining in Eq. (4) the number of photogenerated carriers is obtained from the photoconductivity data, see the dashed line in Fig. 10. The free carrier densities are calculated from the magnetotransport measurements, and χ_c from the data at low intensities where the nonlinearity does not play a role, see squares in Fig. 10, and the hole photocurrent is absent. The larger value of the inversion intensity for the linear polarization, $I_{\text{inv}}^{\text{L}} > I_{\text{inv}}^{\text{C}}$, compare Figs. 3(a) and 3(b), is attributed to the inequality

$$\left|\frac{\chi_v^{(0)}}{\chi_c^{(0)}}\right| > \left|\frac{\chi_v^{(0)} + \chi_v^{(1)}}{\chi_c^{(0)} + \chi_c^{(1)}}\right|.$$

The possibility of the valence band contribution to be indeed larger than that of the conduction band is also supported by numerical computation carried out for a phonon-involved indirect optical transitions. Calculations of the shift photocurrent for free charged carriers described by using the Bernevig–Hughes–Zhang model [12] with indirect optical transition processes involving scattering by acoustic phonons in the (013)-oriented HgTe QWs [49], show that the hole photocurrent can exceed the electron contribution by an order of magnitude. Details of this calculations are out of focus of the current paper.

Finally, it is important to note that a difference between the inversion points for the current measured in the x and diagonal directions, Fig. 10 and the inset in the figure, is caused by the nonparallel directions of the electron and hole photocurrents predicted by Eqs. (20). In agreement with the prediction for the C_1 symmetry, the interplay of the differently oriented vectors $\boldsymbol{J}^{(c)}$ and $\boldsymbol{J}^{(v)}$ yields different inversion points for different pairs of contacts. A larger value of the inversion intensity I_{inv} for the diagonal contacts is attributed to a smaller ratio $|\chi_v/\chi_c|$ in the y direction as compared with that in the x direction. At last but not least we note that the observed shift of the inversion intensity I_{inv} to higher intensity due to temperature increase, see inset in Fig. 4, is in agreement with the reduction of the nonlinearity confirmed by the photoconductivity data, see inset in Fig. 6.

V. SUMMARY

To summarize, our experiments demonstrate that excitation of doped HgTe-based QW structures exhibiting a noninverted band order results in light impact ionization, despite the fact that the Fermi energy exceeds the energy of the forbidden gap. This effect results in strongly nonlinear photogalvanic current exhibiting a dynamic sign inversion and positive photoconductivity due to the electron-pair generation. We show that for a fixed radiation electric field the ionization probability, being proportional to $\exp(-E_0^2/E^2)$ with characteristic field $E_0 \propto \omega$, exponentially increases with the reduction of the square of radiation frequency and/or the Fermi energy level. The probability also rises under sample cooling. The

developed theory of the observed phenomena describes well all characteristic features of the nonlinear optoelectronic phenomena. It demonstrates that under the condition $\omega\tau \gg 1$ the ionization occurs because of collisions in the presence of the oscillating electric field of light wave. A distinguish feature of the *light* impact ionization reported in the present work is that it occurs even for the Fermi level lying in the conduction band. In the latter case, in contrast to all previous works on impact ionization, the electron gas heating is needed to deplete the occupied states enabling the ionization process.

The observed strong nonlinearity, being caused by electron gas heating and electron scattering, reveals importance of scattering processes in the search on Floquet states and realization of Floquet topological insulator. Note that the influence of scattering and electron gas heating effects on the formation of dressed states and band modification are ignored in almost all theoretical works on this topic. As for the future work, we expect that experiments on QW structures with Fermi energy in the forbidden gap should allow to examine photoreponse under conditions with suppressed impact ionization process and may explore the Floquet states formation. Such experiments require structures with a semitransparent gate robust to high power terahertz radiation and allowing application of substantially higher negative gate voltages as used in the present work.

APPENDIX

1. Electron-optical phonon collision integral

We start with description of the loss term of collision integral corresponding to electron transitions out of the state with momentum $\boldsymbol{p}$ due to spontaneous phonon emission. We denote as $\boldsymbol{k} = \boldsymbol{p}/\hbar$ and $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{q}_\parallel$, correspondingly, the electron wave vectors before and after emission of phonon of wave vector $\boldsymbol{q} = \boldsymbol{q}_\parallel + \boldsymbol{q}_z$, where $\boldsymbol{q}_\parallel$ and $\boldsymbol{q}_z$ are the in-plane and z components of vector $\boldsymbol{q}$, respectively (here the momentum conservation law is taken into account). In these terms,

$$St_-^{\mathrm{ph}}\{f\} = \frac{2\pi}{\hbar}\int |M_{\boldsymbol{q}}|^2 \delta(\varepsilon_k - \varepsilon_{k'} - \varepsilon_0)$$
$$\times \theta(\varepsilon_k - \varepsilon_0)f(\boldsymbol{k},t)[1 - f(\boldsymbol{k}',t)]\frac{d^2k'dq_z}{(2\pi)^3}, \quad \text{(A1)}$$

where $M_{\boldsymbol{q}}$ is the matrix element of electron-phonon interaction [50]:

$$M_{\boldsymbol{q}} = \sqrt{\frac{4\pi e^2 \eta \omega_0}{\bar{\epsilon}}} \frac{J(q_z)}{\sqrt{q_\parallel^2 + q_z^2}},$$

$$J(q_z) = \int \psi^2(z)\mathrm{e}^{\mathrm{i}q_z z}dz. \quad \text{(A2)}$$

Here, the Heaviside step function $\theta(\varepsilon - \varepsilon_0) = 1$ at $\varepsilon \geqslant \varepsilon_0$ and $\theta(\varepsilon - \varepsilon_0) = 0$ at $\varepsilon < \varepsilon_0$.

The 2D electron wave function $\psi(z)$ decays on a scale of the quantum well width W. Substituting Eq. (A2) in Eq. (A1) and integrating over q_z and over the modulus of $\boldsymbol{k}'$ for $q_\parallel W \ll 1$ (only the lowest energy level is populated) yield

$$St_-^{\mathrm{ph}}\{f\} = \frac{2\pi^2 e^2 \varepsilon_0 g(\varepsilon - \varepsilon_0)}{\bar{\epsilon}} \int_0^{2\pi} \frac{f(\varepsilon,\varphi,t)[1 - f(\varepsilon - \varepsilon_0, \varphi',t)]\theta(\varepsilon - \varepsilon_0)}{\sqrt{p^2(\varepsilon) + p^2(\varepsilon - \varepsilon_0) - 2p(\varepsilon)p(\varepsilon - \varepsilon_0)\cos(\varphi - \varphi')}} \frac{d\varphi'}{2\pi}. \quad \text{(A3)}$$

Similarly we obtain

$$St_+^{\mathrm{ph}}\{f\} = \frac{2\pi^2 e^2 \varepsilon_0 g(\varepsilon + \varepsilon_0)}{\bar{\epsilon}} \int_0^{2\pi} \frac{[1 - f(\varepsilon,\varphi,t)]f(\varepsilon + \varepsilon_0, \varphi',t)}{\sqrt{p^2(\varepsilon) + p^2(\varepsilon + \varepsilon_0) - 2p(\varepsilon)p(\varepsilon + \varepsilon_0)\cos(\varphi - \varphi')}} \frac{d\varphi'}{2\pi}. \quad \text{(A4)}$$

2. Derivation of balance equations

We look for a solution of the kinetic equation (5) in the form of series

$$f(\varepsilon,\varphi,t) = \sum_{n=0}^{\infty} f_n(\varepsilon,t)\cos(n\varphi). \quad \text{(A5)}$$

Substituting this series in Eq. (5), we obtain an infinite chain of equations for the functions $f_n(\varepsilon,t)$. For high rate of collisions with impurities, the distribution function is almost isotropic, which allows us to consider only the first two terms of the series (A5). The corresponding equations for the functions $f_0(\varepsilon,t)$ and $f_1(\varepsilon,t) \ll f_0(\varepsilon,t)$ have the form:

$$\frac{\partial f_0}{\partial t} + \frac{eEv(\varepsilon)\cos\omega t}{2}\frac{\partial f_1}{\partial \varepsilon} + \frac{eE\cos\omega t}{2p(\varepsilon)}f_1 = -St_-^{\mathrm{ph}}\{f_0\} + St_+^{\mathrm{ph}}\{f_0\}, \quad \frac{\partial f_1}{\partial t} + eEv(\varepsilon)\cos\omega t\frac{\partial f_0}{\partial \varepsilon} = -\frac{1}{\tau_i(\varepsilon)}f_1. \quad \text{(A6)}$$

In accordance with the assumptions made, in the second of these equations we neglect the contribution of phonon scattering to relaxation of the first harmonic, i.e., to the transport relaxation time. At frequencies exceeding all inverse scattering times,

the contribution of the terms oscillating with time to the function $f_0(\varepsilon, t)$ is small, i.e., $f_0(\varepsilon, t) \approx f_0(\varepsilon)$. Therefore, from the second equation it follows that $f_1(\varepsilon, t) = f_c(\varepsilon) \cos \omega t + f_s(\varepsilon) \sin \omega t$. Substituting these functions in Eq. (A6), after simple transformations we obtain equations for $f_0(\varepsilon)$ in the form

$$-\frac{1}{g(\varepsilon)} \frac{\partial}{\partial \varepsilon} \left[g(\varepsilon) D(\varepsilon) \frac{\partial f_0}{\partial \varepsilon} \right] = \frac{1}{g(\varepsilon)} [F(\varepsilon + \varepsilon_0) - F(\varepsilon)], \quad \varepsilon \geqslant \varepsilon_0, \tag{A7}$$

$$-\frac{1}{g(\varepsilon)} \frac{\partial}{\partial \varepsilon} \left[g(\varepsilon) D(\varepsilon) \frac{\partial f_0}{\partial \varepsilon} \right] = \frac{1}{g(\varepsilon)} F(\varepsilon + \varepsilon_0), \quad \varepsilon < \varepsilon_0. \tag{A8}$$

Here

$$F(\varepsilon) = \frac{4\pi e^2 \varepsilon_0 g(\varepsilon - \varepsilon_0) g(\varepsilon)}{p(\varepsilon)} \mathbf{K}\left(\frac{p(\varepsilon - \varepsilon_0)}{p(\varepsilon)} \right) f_0(\varepsilon)[1 - f_0(\varepsilon - \varepsilon_0)], \tag{A9}$$

$\mathbf{K}(x)$ is the complete elliptic integral of the first kind, which equals $\pi/2$ at $x = 0$ and slowly increases with increasing x. Below we will replace it with $\pi/2$. Assuming now that the distribution function changes only slightly on a scale of the order of the phonon energy, we expand the right-hand side of Eq. (A7) in ε_0, and keep the first nonvanishing term only. This gives

$$-\frac{1}{g(\varepsilon)} \frac{d}{d\varepsilon} [g(\varepsilon) \mathcal{J}(\varepsilon)] = 0, \quad \mathcal{J}(\varepsilon) = D(\varepsilon) \frac{\partial f_0}{\partial \varepsilon} + \frac{\varepsilon_0}{\tau_{\mathrm{ph}}(\varepsilon)} f_0(\varepsilon)[1 - f_0(\varepsilon)], \quad D(\varepsilon) = \frac{e^2 E^2 v^2(\varepsilon)}{4\omega^2 \tau_i(\varepsilon)}, \quad \frac{1}{\tau_{\mathrm{ph}}(\varepsilon)} = \frac{2\pi^2 \varepsilon_0 e^2 g(\varepsilon)}{\bar{\epsilon} p(\varepsilon)}. \tag{A10}$$

The quantity $\mathcal{J}(\varepsilon)$ has the meaning of particle flux in the energy space, with the first term describing the diffusive heating of the electron gas, and the second term—the relaxation of energy due to phonon emission. From the condition of vanishing total flux, one obtains Eq. (4) of the main text. Further, Eq. (6) of the main text is obtained from Eq. (A8), where one needs to put

$$F(\varepsilon) = \frac{2\pi^2 e^2 \varepsilon_0 g(\varepsilon - \varepsilon_0) g(\varepsilon)}{p(\varepsilon)} f_0(\varepsilon)[1 - f_0(\varepsilon - \varepsilon_0)], \tag{A11}$$

so that

$$\frac{1}{\tau_{\mathrm{ph}}^+(\varepsilon)} = \frac{2\pi^2 e^2 \varepsilon_0 g(\varepsilon - \varepsilon_0) g(\varepsilon)}{\bar{\epsilon} p(\varepsilon)}. \tag{A12}$$

---

[1] S. D. Ganichev and W. Prettl, *Intense Terahertz Excitation of Semiconductors* (Oxford University Press, Oxford, 2005).

[2] T. Kampfrath, A. Sell, G. Klatt, A. Pashkin, S. Mährlein, T. Dekorsy, M. Wolf, M. Fiebig, A. Leitenstorfer, and R. Huber, Nat. Photon. **5**, 31 (2010).

[3] T. Kampfrath, K. Tanaka, and K. A. Nelson, Nat. Photon. **7**, 680 (2013).

[4] F. Langer, M. Hohenleutner, C. P. Schmid, C. Poellmann, P. Nagler, T. Korn, C. Schüller, M. S. Sherwin, U. Huttner, J. T. Steiner, S. W. Koch, M. Kira, and R. Huber, Nature (London) **533**, 225 (2016).

[5] S. Baierl, M. Hohenleutner, T. Kampfrath, A. K. Zvezdin, A. V. Kimel, R. Huber, and R. V. Mikhaylovskiy, Nat. Photon. **10**, 715 (2016).

[6] H. B. Banks, Q. Wu, D. C. Valovcin, S. Mack, A. C. Gossard, L. Pfeiffer, R.-B. Liu, and M. S. Sherwin, Phys. Rev. X **7**, 041042 (2017).

[7] B. Piętka, N. Bobrovska, D. Stephan, M. Teich, M. Król, S. Winnerl, A. Pashkin, R. Mirek, K. Lekenta, F. Morier-Genoud, H. Schneider, B. Deveaud, M. Helm, M. Matuszewski, and J. Szczytko, Phys. Rev. Lett. **119**, 077403 (2017).

[8] H. A. Hafez, S. Kovalev, J.-C. Deinert, Z. Mics, B. Green, N. Awari, M. Chen, S. Germanskiy, U. Lehnert, J. Teichert, Z. Wang, K.-J. Tielrooij, Z. Liu, Z. Chen, A. Narita, K. Müllen, M. Bonn, M. Gensch, and D. Turchinovich, Nature (London) **561**, 507 (2018).

[9] M. A. W. van Loon, N. Stavrias, N. H. Le, K. L. Litvinenko, P. T. Greenland, C. R. Pidgeon, K. Saeedi, B. Redlich, G. Aeppli, and B. N. Murdin, Nat. Photon. **12**, 179 (2018).

[10] N. H. Lindner, G. Refael, and V. Galitski, Nat. Phys. **7**, 490 (2011).

[11] K.-M. Dantscher, D. A. Kozlov, M. T. Scherr, S. Gebert, J. Bärenfänger, M. V. Durnev, S. A. Tarasenko, V. V. Bel'kov, N. N. Mikhailov, S. A. Dvoretsky, Z. D. Kvon, J. Ziegler, D. Weiss, and S. D. Ganichev, Phys. Rev. B **95**, 201103 (2017).

[12] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[13] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).

[14] S. D. Ganichev, A. P. Dmitriev, S. A. Emel'yanov, Y. V. Terent'ev, I. D. Yaroshetskii, and I. N. Yassievich, Zh. Eksp. Teor. Fiz. **90**, 445 (1986) [JETP **63**, 256 (1986)].

[15] L. V. Keldysh, Zh. Eksp. Teor. Fiz. **48**, 1692 (1965) [JETP **21**, 1135 (1965)].

[16] S. D. Ganichev, J. Diener, and W. Prettl, Appl. Phys. Lett. **64**, 1977 (1994).

[17] A. G. Markelz, N. G. Asmar, B. Brar, and E. G. Gwinn, Appl. Phys. Lett. **69**, 3975 (1996).

[18] P. Gaal, K. Reimann, M. Woerner, T. Elsaesser, R. Hey, and K. H. Ploog, Phys. Rev. Lett. **96**, 187402 (2006).

[19] H. Wen, M. Wiczer, and A. M. Lindenberg, Phys. Rev. B **78**, 125203 (2008).

[20] M. C. Hoffmann, J. Hebling, H. Y. Hwang, K.-L. Yeh, and K. A. Nelson, J. Opt. Soc. Am. B **26**, A29 (2009).

[21] H. A. Hafez, X. Chai, A. Ibrahim, S. Mondal, D. Férachou, X. Ropagnol, and T. Ozaki, J. Opt. **18**, 093004 (2016).

[22] A. T. Tarekegne, H. Hirori, K. Tanaka, K. Iwaszczuk, and P. U. Jepsen, New J. Phys. **19**, 123018 (2017).

[23] I. V. Oladyshkin, S. B. Bodrov, Y. A. Sergeev, A. I. Korytin, M. D. Tokman, and A. N. Stepanov, Phys. Rev. B **96**, 155401 (2017).

[24] S. A. Dvoretsky, N. N. Mikhailov, Y. G. Sidorov, V. A. Shvets, S. N. Danilov, B. Wittman, and S. D. Ganichev, J. Electron. Mater. **39**, 918 (2010).

[25] H. Maier, J. Ziegler, R. Fischer, D. Kozlov, Z. D. Kvon, N. Mikhailov, S. A. Dvoretsky, and D. Weiss, Nat. Commun. **8**, 2023 (2017).

[26] C. Drexler, N. Dyakonova, P. Olbrich, J. Karch, M. Schafberger, K. Karpierz, Y. Mityagin, M. B. Lifshits, F. Teppe, O. Klimenko, Y. M. Meziani, W. Knap, and S. D. Ganichev, J. Appl. Phys. **111**, 124504 (2012).

[27] P. Olbrich, J. Karch, E. L. Ivchenko, J. Kamann, B. März, M. Fehrenbacher, D. Weiss, and S. D. Ganichev, Phys. Rev. B **83**, 165320 (2011).

[28] P. Olbrich, L. E. Golub, T. Herrmann, S. N. Danilov, H. Plank, V. V. Bel'kov, G. Mussler, C. Weyrich, C. M. Schneider, J. Kampmeier, D. Grützmacher, L. Plucinski, M. Eschbach, and S. D. Ganichev, Phys. Rev. Lett. **113**, 096601 (2014).

[29] S. D. Ganichev, S. A. Emel'yanov, and I. D. Yaroshetskii, Pis'ma Zh. Eksp. Teor. Fiz. **35**, 297 (1982) [JETP Lett. **35**, 368 (1982)].

[30] S. D. Ganichev, Y. V. Terent'ev, and I. D. Yaroshetskii, Pis'ma Zh. Tekh. Fiz. **11**, 46 (1985) [Sov. Tech. Phys. Lett. **11**, 20 (1989)].

[31] S. Ganichev, Physica B **273-274**, 737 (1999).

[32] V. V. Bel'kov, S. D. Ganichev, E. L. Ivchenko, S. A. Tarasenko, W. Weber, S. Giglberger, M. Olteanu, H. P. Tranitz, S. N. Danilov, P. Schneider, W. Wegscheider, D. Weiss, and W. Prettl, J. Phys. Cond. Matt. **17**, 3405 (2005).

[33] B. Wittmann, S. N. Danilov, V. V. Bel'kov, S. A. Tarasenko, E. G. Novik, H. Buhmann, C. Brüne, L. W. Molenkamp, Z. D. Kvon, N. N. Mikhailov, S. A. Dvoretsky, N. Q. Vinh, A. F. G. van der Meer, B. Murdin, and S. D. Ganichev, Semicond. Sci. Technol. **25**, 095005 (2010).

[34] S. V. Morozov, M. S. Joludev, A. V. Antonov, V. V. Rumyantsev, V. I. Gavrilenko, V. Y. Aleshkin, A. A. Dubinov, N. N. Mikhailov, S. A. Dvoretskiy, O. Drachenko, S. Winnerl, H. Schneider, and M. Helm, Fiz. Tekh. Poluprovodn. **46**, 1388 (2012) [Semicond. **46**, 1362 (2012)].

[35] S. D. Ganichev, A. P. Dmitriev, S. A. Emel'yanov, Y. V. Terent'ev, I. D. Yaroshetskii, and I. N. Yassievich, Pis'ma Zh. Eksp. Teor. Fiz. **40**, 187 (1984) [JETP Lett. **40**, 948 (1984)].

[36] S. H. Groves, R. N. Brown, and C. R. Pidgeon, Phys. Rev. **161**, 779 (1967).

[37] L. C. Lew Yan Voon and M. Willatzen, *The $k \cdot p$ Method* (Springer Berlin, Heidelberg, 2009).

[38] E. G. Novik, A. Pfeuffer-Jeschke, T. Jungwirth, V. Latussek, C. R. Becker, G. Landwehr, H. Buhmann, and L. W. Molenkamp, Phys. Rev. B **72**, 035321 (2005).

[39] V. F. Gantmakher and Y. B. Levinson, *Carrier Scattering in Metals and Semiconductors* (North Holland, Amsterdam, 1988).

[40] Note that the matrix element of electron scattering by short range impurities does not depend on energy. Thus energy dependence of the scattering rate, being proportional to the density of states (DOS) is solely determined by the energy dependence of the DOS.

[41] V. I. Belinicher, E. L. Ivchenko, and B. I. Sturman, Zh. Eksp. Teor. Fiz. **83**, 649 (1982) [JETP **56**, 359 (1982)].

[42] B. I. Sturman and V. M. Fridkin, *The Photovoltaic and Photorefractive Effects in Non-Centrosymmetric Materials* (Gordon and Breach Science Publishers, New York, 1992).

[43] J. E. Sipe and A. I. Shkrebtii, Phys. Rev. B **61**, 5337 (2000).

[44] E. L. Ivchenko, *Optical Spectroscopy of Semiconductor Nanostructures* (Alpha Sci. Int., Harrow, UK, 2005).

[45] L. E. Golub and E. L. Ivchenko, Zh. Eksp. Teor. Fiz. **139**, 175 (2011) [JETP **112**, 152 (2011)].

[46] S. M. Young and A. M. Rappe, Phys. Rev. Lett. **109**, 116601 (2012).

[47] U. Bajpai, B. S. Popescu, P. Plechac, B. K. Nikolic, L. E. F. Foa Torres, H. Ishizuka, and N. Nagaosa, arXiv:1803.04404.

[48] V. I. Belinicher and B. I. Sturman, Usp. Fiz. Nauk **130**, 415 (1980) [Sov. Phys. Usp. **23**, 199 (1980)].

[49] P. Olbrich, C. Zoth, P. Vierling, K.-M. Dantscher, G. V. Budkin, S. A. Tarasenko, V. V. Bel'kov, D. A. Kozlov, Z. D. Kvon, N. N. Mikhailov, S. A. Dvoretsky, and S. D. Ganichev, Phys. Rev. B **87**, 235439 (2013).

[50] A. I. Anselm, *Introduction to Semiconductor Theory* (Prentice Hall, New Jersey, 1981).