State-resolved ultrafast dynamics of impact ionization in InSb

H. Tanimura, J. Kanasaki & K. Tanimura

The Institute of Scientific and Industrial Research, Osaka University, 8-1, Mihogaoka, Ibaraki, Osaka 567-0047, Japan.

Impact ionization (IMP) is a fundamental process in semiconductors, which results in carrier multiplication through the decay of a hot electron into a low-energy state while generating an electron-hole pair. IMP is essentially a state selective process, which is triggered by electron-electron interaction involving four electronic states specified precisely by energy and momentum conservations. However, important state-selective features remain undetermined due to methodological limitations in identifying the energy and momentum of the states involved, at sufficient temporal resolution, to reveal the fundamental dynamics. Here we report state-resolved ultrafast hot electron dynamics of IMP in InSb, a semiconductor with the lowest band-gap energy. The ultrafast decay of state-resolved hot-electron populations and the corresponding population increase at the conduction band minimum are directly captured, and the rate of IMP is unambiguously determined. Our analysis, based on the direct knowledge of state-resolved hot electrons, provides far deeper insight into the physics of ultrafast electron correlation in semiconductors.

IMP is an important process induced by Coulomb interactions among electrons in semiconductors. It critically affects carrier transport under high-electric field conditions, and is of currently studied as a means to improve photovoltaic-device efficiency. Extensive studies have been carried out on IMP in semiconductors using a variety of experimental and theoretical techniques. Traditional experimental studies primarily measure the electric-field dependence of carrier multiplication. Apart from the difficulty in correlating the electric field with the excess energy of carriers, information concerning state-selective characteristics is completely lost by the energy-integrating nature of this method. Additional spectroscopic studies, which could reduce these drawbacks, neither specify the pertinent energy and momentum states nor determine the absolute rate directly. The rate of IMP has been analytically formulated for a simple case, but there have been no experimental results for comparison with those predicted by this formula, leaving our fundamental understanding of the IMP process ambiguous and incomplete. Therefore, it is critical to measure quantum-mechanical dynamics of IMP on a state-resolved basis.

Here we directly capture state-resolved dynamics of IMP in InSb, using angle- and time-resolved photoemission spectroscopy. InSb crystals with (110) surfaces were excited by fs-laser pulses, and photoinjected hot electrons were detected using time-delayed 4.71-eV probe pulses. By detecting photoelectrons as a function of energy and emission angle, we directly determine energy- and momentum-resolved distributions of hot electrons at fs-temporal resolution.

Results

Figure 1 (a) shows a photoemission map, at a time delay (Δt) of 50 fs, as a function of ε and θ. The white broken and solid curves are the energy dispersions of the conduction band along the Γ-X and Γ-L lines. The pump pulse injects electrons at energies of 1.10 eV, 0.76 eV, and 0.40 eV above the conduction-band minimum (CBM) from the heavy-hole (HH), light-hole (LH), and split-off (SO) valence band, respectively. Importantly, the transition from the HH band around L injects hot electrons predominantly into the L valley. The angle- and energy-integrated intensity of the peaks from the HH band shares more than 80% of total photoemission intensity generated by the pump pulse, which creates carriers at a density of 4.2 × 10^{17} cm^{-3}.

Figures 1 (b) and (c) display angle-integrated intensities of photoemissions from hot electrons and electrons near the CBM as a function of pump-probe delay, Δt. Electrons injected from the SO valence band decay within 200 fs after excitation, while those from LH band decay more rapidly within 100 fs. However, electrons injected into Γ and L valleys from the HH band exhibit relatively slow decay, surviving until ~800 fs. Interestingly, a new peak at E_{ex} = 0.63 eV begins to appear at around Δt = 400 fs. The most striking feature is that the population near the CBM increases dramatically at Δt < 300 fs. As seen in Fig. 1 (b), the electrons injected from the HH band
We can distinguish two distinctive processes in the growth of the population near the CBM: the fast process, at $\Delta t < 300$ fs, and the slow process which follows it. During the slow process, the hot-electron population decreases in the L valley, and a new photoemission band appears at $E_{\text{ex}}$ where the new band appears. Therefore, we conclude that the IMP by hot electrons at the L$_1$ valley is phonon-assisted. The broken curve in Fig. 2(a) shows electron population dynamics near the L$_1$ valley at the position, $E(\Gamma_1)$. In contrast to the slow process, the fast process is temperature insensitive. A stepwise increase in the population near the CBM is clearly detected at 90 K with almost the same enhancement as observed at 293 K. Considering this result along with the ultrafast temporal response, we conclude that the fast process is due to the direct IMP by hot electrons in the $\Gamma$ valley. In fact, electrons injected from the SO and LH bands decay within this temporal domain as shown in Fig. 2(b). The first feature represents the energy relaxation towards the L valley (L$_1$) minimum. We note that photoemissions beyond $\theta = 35^\circ$ cannot be detected in our detection window, resulting in the observed intensity loss. A recent GW band calculation predicts L$_1$ at 0.61 eV above the CBM, almost the same energy as $E_{\text{ex}}$ where the new band appears. Therefore, we conclude that the new band’s appearance is due to the electrons scattered from L$_1$ to the $\Gamma$ valley at the position, $E(\Gamma_1)$. In fact, the rate equation model successfully correlates the relaxation in the L valley, the accumulation at L$_1$, and the L-to-$\Gamma$ intervalley scattering, and satisfactorily describes the population dynamics at $E(\Gamma_1)$ (see solid curves in Fig. 2(b)). The broken curve shows the predicted changes in the population at the minimum of the L valley.

Importantly, the L-to-$\Gamma$ intervalley scattering and the increase in population near the CBM takes place in parallel as shown in Fig. 2(c); both peaks near the CBM and electrons at $E(\Gamma_1)$ are increasing concurrently. These results reveal that the hot electrons at the L$_1$ induce both IMP and intervalley scattering competitively. In this IMP process, however, the momentum conservation among electronic states can no longer be satisfied, as no valence states exist on the line connecting L$_1$ and the CBM, as shown by red broken line in Fig. 2(d). Therefore, IMP by hot electrons at L$_1$ is phonon-assisted. The red curve in Fig. 2(a) shows electron population dynamics near the CBM measured at 90 K under the same excitation condition as that measured at 293 K. The enhancement in the slow process is dramatically suppressed at 90 K relative to that at 293 K, supporting our conclusion that the IMP by hot electrons at the L$_1$ is phonon-assisted.

In contrast to the slow process, the fast process is temperature insensitive. A stepwise increase in the population near the CBM is clearly detected at 90 K with almost the same enhancement as observed at 293 K. Considering this result along with the ultrafast temporal response, we conclude that the fast process is due to the direct IMP by hot electrons in the $\Gamma$ valley. In fact, electrons injected from the SO and LH bands decay within this temporal domain as may be discerned from Fig. 1(b), with decay times of $\tau_d = 150$ and 50 fs, respectively. In the direct IMP of hot electrons in the $\Gamma$ valley, interaction with the LH valence band is crucial. As seen in Fig. 2(d), hot electrons in the $\Gamma$ valley are under ideal conditions to satisfy required energy and momentum conservations because of the approximately linear dispersions of the conduction band and the LH valence band.

Since the electrons from LH band are injected above the energy of L$_1$, they can be scattered into the L valley. To determine the effects of this intervalley scattering on $\tau_d$, we excited InSb with light pulses at photon energies ranging from 1.2 to 1.57 eV, and determined the $\tau_d$’s of wave packets formed by photoionization. Typical results are shown in Fig. 3(a). The electrons with $E_{\text{ex}} > 0.53$ eV show $\tau_d = 150$ fs, while those with $E_{\text{ex}} > 0.65$ eV show shorter $\tau_d$; the decay time of electrons...
Figure 2 | Temporal changes in the population near the CBM and two processes of impact ionization. (a) Temporal changes in the electron population near the CBM measured at 90 K (red line) and 293 K (black line). Angle and energy integrated photoemission intensities (from −5° to 5° and 0.1 to 0.35 eV) are plotted as a function of Δt. The magnitude of the integrated intensity was converted to the electron density based on spectral-shape analysis (see the text). (b) Temporal changes in the intensities of hot electrons in the L valley (blue line) and those populated at 0.35 eV) are plotted as a function of m. The same scaling factor to convert the integrated photoemission intensity into the electron density in (a) was assumed for these hot electrons. The thin solid curves are the results of the rate-equation analysis, showing temporal populations of relaxing electrons in the L valley and at Γ(L). The broken curve is the result obtained by the analysis, predicting temporal change in the population at the L1. (c) Angle integrated photoemission spectra (−5° to 5°) measured at Δt = 1.2 and 2.5 ps. The two peaks at 0.07 and 0.63 eV represent populations near the CBM and at the (Γ1), respectively. (d) Dispersions along Γ-L direction for valence and conduction bands (after Ref. 18). The black arrow shows the optical transition from LH to the conduction band, and pathways of intervalley scattering from Γ-to-L and L-to-Γ are shown by blue arrows. A possible direct impact ionization process is shown by red arrows. The red broken line connecting the L1 and the CBM represents the limit of the direct impact ionization process.

with $E_{\text{ex}} = 0.79$ eV as is short as 50 fs. The relation between the transition rate, given by the inverse of $\tau_b$, and $E_{\text{ex}}$ is plotted in Fig. 3(b). It is clear that there is a critical energy of $E_{\text{ex}} = 0.63$ eV, exactly coincides with the energy of L1, above which $\tau_b$ becomes dramatically shorter. Therefore, hot electrons with $E_{\text{ex}} > 0.65$ eV undergo efficient scattering to the L valley, while those with $E_{\text{ex}} < 0.60$ eV do not.

Discussion
The results described above show that the hot electrons with $E_{\text{ex}} = 0.5$ eV decay with $\tau_b = 150$ fs. Radiative and phonon-assisted non-radiative recombination occurs at much slower rates\textsuperscript{23}. The possible transient Auger recombination of the hot electrons, induced by photoinjected holes in the LH band, may be ignored as the excitation density of the LH band is at most $1.54 \times 10^{17}$ cm$^{-3}$, which is much smaller than the density of states ($>1.0 \times 10^{18}$ cm$^{-3}$) of the LH band in the energy range where holes are injected. In fact, the decay times of hot electrons with $E_{\text{ex}} < 0.6$ eV did not depend on the excitation density between $2.2 \times 10^{17}$ and $1.3 \times 10^{18}$ cm$^{-3}$. Therefore, we conclude that the decay of these hot electrons is due to direct IMP. Furthermore, as shown in Fig. 3(b), we can determine directly the rate of direct IMP of electrons with $E_{\text{ex}} = 0.5$ eV in InSb to be $1/\tau_b = 6.7 \times 10^{12}$ s$^{-1}$.

The theory of IMP provides the following formula for the rate of IMP for a parabolic system:

$$W_{\text{IMP}} = W_0 \left( \frac{\varepsilon_0}{E_{\text{ex}}} \right)^2 \frac{m_i^*}{m_0} \frac{I_1^2}{I_0} \frac{1}{(1 + 2\mu)^{3/2}} \left( \frac{E_{\text{ex}} - E_T}{E_T} \right)^2,$$

where $W_0 = 4.14 \times 10^{16}$ s$^{-1}$, $\varepsilon_0$ is the free-space permittivity, $E_{\text{ex}}$ is the high-frequency permittivity, $m_i^*$ is the electron effective mass, $\mu = m_i^*/m_e$, $m_e$ is the hole effective mass, $I_1$ and $I_0$ are the cell-periodic overlap integrals, $E_T$ is the threshold energy and $E_T$ is the band-gap energy\textsuperscript{17}. In a parabolic system, $E_T$ is given by $E_T = E_g(1 + 2\mu)/(1 + \mu)$. Using $m_i^*$ for $H$ valence band in InSb, we estimate $E_T = 0.25$ eV at 293 K. When we apply this formula to hot electrons with $E_{\text{ex}} = 0.5$ eV in InSb, we get $W_{\text{IMP}} = 1.1 \times 10^{12}$ s$^{-1}$ even for $I_1^2/I_0^2 = 1$, establishing an upper bound. Given that (1) neglects any screening effects, this theory underestimates the rate significantly. We presume that the discrepancy comes from the parabolic-band approximation in Eq. (1). As seen in Fig. 2(d), the dispersion of the conduction band and LH band is not parabolic, but can be well approximated by a linear dispersion in the energy range from −0.5 to 1.0 eV, except for the gap region. Dispersions along Γ-Γ and Γ-Γ directions show the same characteristic\textsuperscript{18}. Therefore, momentum and energy conservation can be satisfied for the four quantum states involved in IMP processes, without any limitations to the initial impacting electronic state $E_{\text{ex}} > E_T$. This would make the rate of IMP significantly larger when compared to the parabolic case, as we observe.

In conclusion, we have demonstrated the state-selective features in IMP processes for the first time, and made it possible to compare experimental results and theoretical modeling of IMP rates quantitatively. Our study, based on direct measurements of energy and momentum-resolved electron distributions, provides us deeper insight into the IMP processes in semiconductors than previous studies, and gives a sound basis for investigating the current topic of ultrafast interaction of intense THz waves with matter\textsuperscript{20-24}.

Methods
Non-doped InSb (100) wafers were cleaved under ultrahigh vacuum conditions (<5 × 10$^{-10}$ Torr). Surface structures were characterized in situ by a scanning tunneling microscope prior to photoemission measurements\textsuperscript{12}. The (110) surfaces displayed well-ordered (1 × 1)-structure with linear In and Sb rows, and a surface-defect concentration less than 0.5%. We used two different femtosecond laser systems. A mode-locked Ti:sapphire laser, operated at a 76 MHz repetition rate, was used to generate fs-laser pulses between 820–730 nm. The fundamental and its third har-
Figure 3 | Decay times of hot electrons and relation to excess energy. (a) Temporal changes in the photoemission intensities from hot electrons injected at $E_{ex} = 0.79, 0.65$ and $0.54$ eV (violet, green, and red lines, respectively) probed on a semi-logarithmic scale. The blue curve, labelled CC, shows the cross correlation trace between pump and probe pulses, which defines the zero time delay. Measurements were made at 293 K. The overlap between pump and probe pulses can be neglected at $\Delta t > 150$ fs. Decay times can be determined almost uniquely from the plot, although the analysis for a whole temporal domain including rise time needs more sophisticated methods using the optical Bloch equation (see supplementary information). The thin black solid lines show results of analysis using the optical Bloch equation. (b) The relationship between transition rate, defined by the inverse of the decay time, of hot electrons and excess energy. Green and red dots are for the hot electrons injected from light-hole and split-off valence bands, respectively. The solid curve shows the dispersion of the conduction band along the $\Gamma$–L direction (after Ref. 18). The theoretical position of $L_1$ is 0.61 eV above the CBM, while experimentally determined value is 0.63 eV above the CBM as shown by the broken line.

monic, generated using beta-barium borate crystals, were used as pump and probe pulses. Temporal widths of the pump and the probe pulses were 80 fs and 100 fs at the sample position in the UHV chamber, giving a cross-correlation trace of 125 fs. To study the temporal widths of the pump and the probe pulses were 80 fs and 100 fs at the sample position in the UHV chamber, giving a cross-correlation trace of 125 fs. To study the temporal widths of the pump and the probe pulses were 80 fs and 100 fs at the sample position in the UHV chamber, giving a cross-correlation trace of 125 fs. To study the temporal widths of the pump and the probe pulses were 80 fs and 100 fs at the sample position in the UHV chamber, giving a cross-correlation trace of 125 fs. To study the temporal widths of the pump and the probe pulses were 80 fs and 100 fs at the sample position in the UHV chamber, giving a cross-correlation trace of 125 fs. To study the temporal widths of the pump and the probe pulses were 80 fs and 100 fs at the sample position in the UHV chamber, giving a cross-correlation trace of 125 fs. To study the temporal widths of the pump and the probe pulses were 80 fs and 100 fs at the sample position in the UHV chamber, giving a cross-correlation trace of 125 fs. To study the temporal widths of the pump and the probe pulses were 80 fs and 100 fs at the sample position in the UHV chamber, giving a cross-correlation trace of 125 fs.

The ionization energy of InSb(110) is 4.77 eV 23. Therefore, the third harmonic of the 790-nm fundamental beam (4.71 eV) can probe the states populated in the conduction band, with completely suppressing intense signals due to direct photoionization from the valence band.

For photoemission spectroscopy, a hemispherical electron analyzer, equipped with an angle-resolved lens mode and a two-dimensional image-type detector served as the electron spectrometer. The entrance slit of the analyzer was placed on the optical plane defined by incoming and reflected light. Two-dimensional images of photoelectrons were recorded as functions of energy and of emission angle $\theta$ along the [001] crystallographic direction. Here, $\theta$ corresponds to a surface-parallel wavevector $k_x$ along the $\Gamma$–$Y$ direction of the surface Brillouin zone (SBZ), on which all states along the $\Gamma$–$X$ and $\Gamma$–$L$ lines, in the bulk Brillouin zone, are projected with finite values of $k_y$. The instrumental energy resolution with fs-probe light was 45 meV, while angle resolution was limited in the range of $\pm 1^\circ$.