

Ultrafast hot-carrier-dominated photocurrent in graphene

Dong Sun¹, Grant Aivazian¹, Aaron M. Jones¹, Jason S. Ross², Wang Yao³, David Cobden¹ and Xiaodong Xu^{1,2*}

The combination of its high electron mobility^{1–5}, broadband absorption⁶ and ultrafast luminescence^{7–10} make graphene attractive for optoelectronic and photonic applications^{11–13}, including transparent electrodes¹⁴, mode-locked lasers¹⁵ and high-speed optical modulators¹⁶. Photo-excited carriers that have not cooled to the temperature of the graphene lattice are known as hot carriers, and may limit device speed and energy efficiency. However, their roles in charge and energy transport are not fully understood^{17–20}. Here, we use time-resolved scanning photocurrent microscopy to demonstrate that hot carriers, rather than phonons, dominate energy transport across a tunable graphene p–n junction excited by ultrafast laser pulses. The photocurrent response time varies from 1.5 ps at room temperature to 4 ps at 20 K, implying a fundamental bandwidth of ~500 GHz (refs 12,13,21). Gate-dependent pump-probe measurements demonstrate that both thermoelectric and built-in electric field effects contribute to the photocurrent, with the contribution from each depending on the junction configuration. The photocurrent produced by a single pulsed laser also displays multiple polarity reversals as a function of carrier density, which is a possible signature of impact ionization^{19,22,23}.

The development of tunable, high-speed and broadband graphene photodetectors has attracted interest as a first step towards optoelectronic devices^{12,13,17–19,24–28}. Previous work has demonstrated a detection bandwidth of ~40 GHz for a graphene/metal contact (GM) device¹³, and simple analysis suggests an operation bandwidth of ~500 GHz (ref. 13). On the other hand, despite the successful demonstration of a prototype GM photodetector, basic questions related to the photocurrent-generation mechanism with highly non-equilibrium electrons (the electron temperature is much higher than the lattice temperature) remain, such as whether the built-in field or the photo-thermoelectric effect dominates the photocurrent generation^{12,13,17–19,24,25,27,29}, and the role of hot-carrier transport in ultrafast devices and possible carrier multiplication processes¹⁹. This is mainly due to the speed limitations of electrical measurements and to the use of continuous-wave (c.w.) excitation^{13,16,21}, in which the electron temperature (T_e) is close to the lattice temperature (T_L). For picosecond timescale operation, non-equilibrium effects come into play because the hot carriers do not have time to reach thermal equilibrium with the lattice. Moreover, even under c.w. excitation, due to the relatively slow (~10 ns) acoustic phonon scattering rate, recent theory indicates that hot-carrier transport and carrier multiplication^{19,22} could dominate photocurrent generation. Therefore, it is important to study experimentally hot-carrier transport with sub-picosecond temporal resolution.

We therefore applied spatially (~1 μm) and temporally (~250 fs) resolved scanning photocurrent microscopy (SPM) to investigate tunable graphene p–n junctions. We used a dual-gate structure, which allowed independent control of the charge density in adjacent regions¹⁷. We studied both single- and few-layer exfoliated graphene as well as single-layer graphene grown by chemical vapour deposition (CVD). All demonstrated similar behaviour. Measurements on CVD graphene are presented in the Supplementary Information. We focus here on an exfoliated trilayer device, which was identified as ABA stacked by Raman spectroscopy (Supplementary Fig. S1). Figure 1a presents a schematic diagram of the measurement set-up, and Fig. 1b (left) shows an optical microscope image of the device. Application of voltages to the top and bottom gate creates a graphene homojunction on either side of the top gate that can be tuned between p–p, p–n, n–p and n–n configurations (where the first and second letters represent the doping outside and under the top gate, respectively). Because the fabrication process induces doping, which varies between samples, we define the effective back-gate voltage V_{bg} to be measured relative to the gate voltage neutrality point (+70 V for the trilayer device). Figure 1c shows resistance as a function of V_{bg} and top-gate voltage V_{tg} . The sloping dashed line indicates where the graphene under the top gate is at the neutrality point. We estimate the mobility of this device to be ~200 $\text{cm}^2 \text{V}^{-1} \text{s}^{-1}$. The low mobility in this case is partially due to the deposition of the top gate.

We first performed standard photocurrent microscopy by scanning the focused pulsed laser over the device (Supplementary Fig. S2). The results of such a measurement at a temperature of 20 K and laser power of ~70 μW , or ~4 $\times 10^{14}$ photons per cm^2 per pulse, are shown in Fig. 1b. Photocurrent is observed at the graphene–metal contact interfaces and at the p–n junctions. We focus on the latter. We first consider the results of excitation by a single pulsed laser. Figure 1d shows the photocurrent as a function of gate voltages with the laser spot located at the p–n junction indicated by the circle in Fig. 1b. Figure 1e is a single line trace at $V_{\text{bg}} = 30 \text{ V}$, which corresponds to global n-doping. We observe that the photocurrent crosses zero twice as V_{tg} varies. This is a clear indication of photocurrent generation from the thermoelectric effect. The built-in field effect only changes sign once as a function of V_{tg} , and only the addition of the thermoelectric effect can account for multiple zero crossings¹⁹. We note that Fig. 1b is different from the symmetric sixfold pattern predicted by ref. 19 and observed in ref. 20. The deviation we see is explained by the slightly different mobility of the graphene covered by the top gate and the interplay between built-in electric fields and the photo-thermoelectric effect under pulsed excitation, as will be discussed in detail below.

¹Department of Physics, University of Washington, Seattle, Washington 98195, USA, ²Department of Materials Science and Engineering, University of Washington, Seattle, Washington 98195, USA, ³Department of Physics and Center of Theoretical and Computational Physics, The University of Hong Kong, Hong Kong, China. *e-mail: xuxd@uw.edu

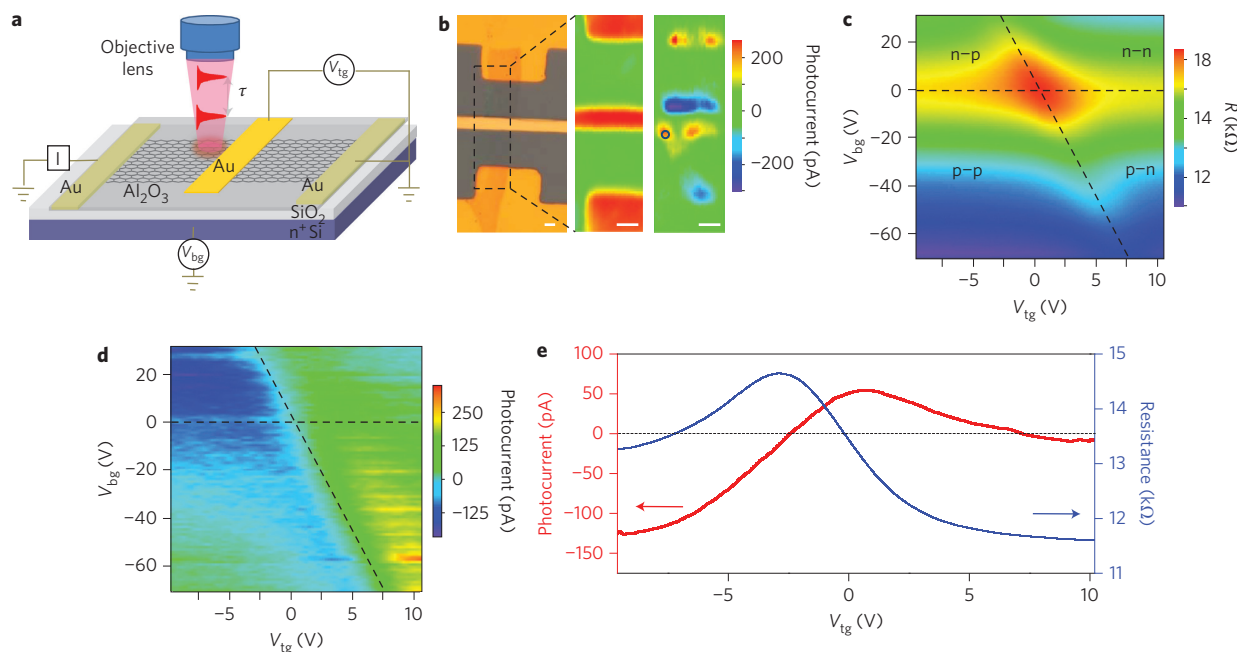


Figure 1 | Standard photocurrent microscopy of a graphene device. **a**, Experimental set-up and schematic structure of a graphene device with top (gold) and bottom (dark grey) gates. **b**, Left: optical microscope image of the graphene trilayer device; centre: scanning reflection image; right: photocurrent image obtained at $V_{tg} = 10$ V, $V_{bg} = 0$ V, and laser power of $70 \mu\text{W}$ at 20 K. Scale bars, $3 \mu\text{m}$. **c**, Source-drain resistance as a function of V_{tg} and V_{bg} . **d**, Photocurrent as a function of V_{tg} and V_{bg} with the laser fixed at the p-n junction (circle in **b**). The diagonal dashed line in **c** and **d** indicates where charge neutrality occurs under the top gate. **e**, Photocurrent (red) and source-drain resistance (blue) as a function of V_{tg} at $V_{bg} = 30$ V.

We next applied ultrafast pump-probe techniques to investigate the hot-carrier dynamics. Two independently controlled pulsed beams were focused at the same spot. The pump and probe pulses were cross-polarized to minimize interference near zero delay.

The probe beam was chopped and a lock-in amplifier used to detect the resulting modulation of the photocurrent. The measured signal can be thought of as the probe-induced photocurrent (here termed PC), a quantity whose sensitivity to the presence of the pump can be investigated as a function of delay and pulse amplitude (Supplementary Figs S3 and S4).

Figure 2a presents typical pump-probe measurements at a 20 K. The three data sets correspond to different pump powers. The measurements have been normalized to the PC generated in the absence of the pump, and all pump-probe data are for a probe power of $70 \mu\text{W}$ unless otherwise specified. The presence of the pump only affects the PC near zero delay, producing a sharp dip with the appearance of a saturation of PC generation with increasing excitation power. The response time τ (defined as the half-width at half-maximum of the dip) is ~ 4 ps, with no clear dependence on pump power (red symbols in Fig. 2b). The PC at zero delay is plotted as a function of pump power in Fig. 2b (black symbols). The decrease of PC as pump power increases is also consistent with a saturation effect, as is the saturation of the PC as a function of probe power with no pump, as shown in Fig. 2c. The latter has an approximate dependence $I_{pc} \approx P^{0.7}$ on probe power P .

The dynamics change significantly with environmental temperature, as can be seen in Fig. 3a, which shows measurements at 250 K and 20 K. We find that τ increases from ~ 1.5 ps to ~ 4 ps on cooling from 295 K to 20 K (Fig. 3b). Interestingly, over the same temperature range the PC generated by c.w. excitation increases by a factor of ~ 13 while the PC generated by pulsed excitation at the same power is almost unchanged (Fig. 3c and Supplementary Fig. S5). Similar results were obtained with single-layer graphene, both exfoliated and CVD-grown, and for mobilities ranging from $\sim 200 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ to $\sim 2,000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ (Supplementary Figs S6 and S7). As shown in Fig. 3d, there is no clear variation of the PC dynamics between these cases.

We now discuss the implications of these findings. It is known that the pulse-excited electrons in graphene equilibrate among

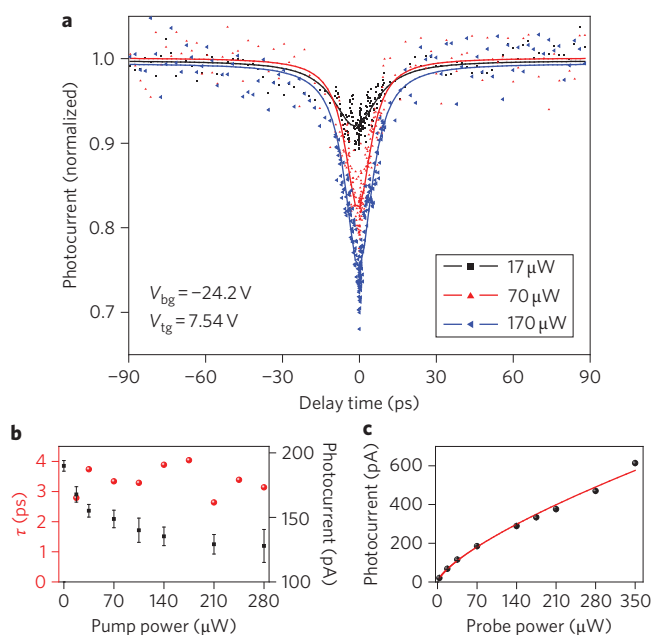


Figure 2 | Power dependence of the pump-probe measurements. **a**, Probe-induced photocurrent at a p-n junction as a function of pump-probe pulse delay. Probe power, $70 \mu\text{W}$; pump power as indicated; $V_{tg} = 7.5$ V and $V_{bg} = -24.2$ V. Solid lines are guides to the eye. **b**, Probe-induced photocurrent at zero delay (black) and response time τ (red), defined as the half-width at half-maximum of the dip in **a** as a function of pump power. **c**, Photocurrent as a function of probe power P with no pump. The line is a power-law fit with $I_{pc} \approx P^{0.7}$.

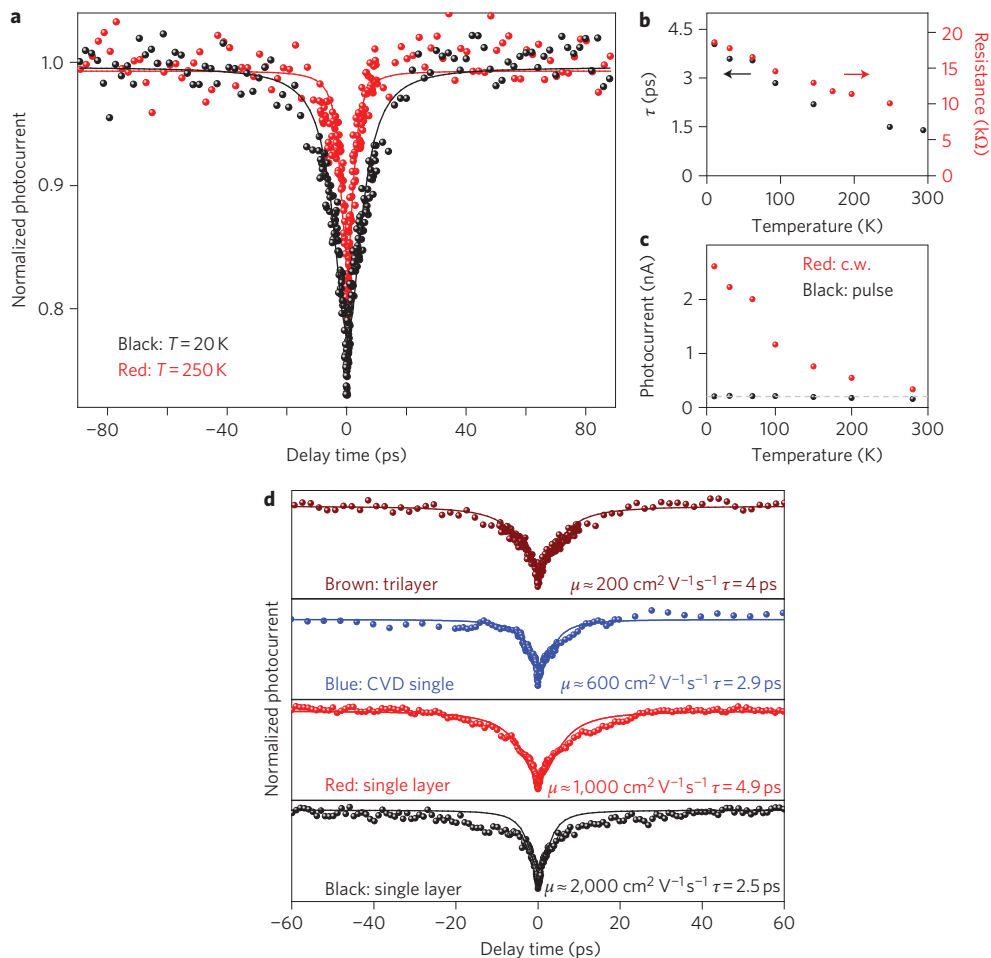


Figure 3 | Temperature dependence of photocurrent amplitude and dynamics. **a**, Delay time scan of probe-induced photocurrent at 250 K (red) and 20 K (black). Probe power, 70 μW ; pump power, 270 μW ; $V_{\text{tg}} = 7.5$ V and $V_{\text{bg}} = -24.2$ V. **b**, Temperature dependence of source-drain resistance (red) and response time τ (black). **c**, Temperature dependence of photocurrent generated by c.w. (red) and 250 fs pulsed (black) excitation at 70 μW . **d**, Mobility dependence of PC dynamics. Time delay scan of PC at 20 K on exfoliated single layers (black and red) with different mobilities, a CVD-grown single layer (blue) and a graphene trilayer (brown). Probe power, 70 μW ; pump power, 140 μW .

themselves on a timescale of tens of femtoseconds through electron–electron interactions, after which the electron system can be characterized by a local electron temperature^{9,30,31}. This temperature, T_e , determines the PC generated by both the built-in field and photo-thermoelectric effects. Understanding the details of the observed saturation effect requires a knowledge of the specific heats of electrons and optical phonons and of the thermoelectric coefficient at elevated T_e , which is not available at present. Roughly speaking, however, saturation implies that the PC increases more slowly with T_e than T_e increases with laser power; that is, if $I_{\text{pc}} \approx T_e^\alpha$ and $T_e \approx P^{1/\beta}$ so that $I_{\text{pc}} \approx P^{\alpha/\beta}$, then $\alpha < \beta$ (Supplementary Fig. S8).

Pulsed excitation is significantly different in nature from c.w. excitation. The c.w. excitation results in a steady, low T_e (refs 1,25). Under pulsed excitation, T_e has a spike following each pulse, corresponding to the generation of hot carriers, which rapidly cool long before the next pulse arrives. This cooling involves two mechanisms. The first of these is optical phonon-mediated cooling, which reduces T_e to ~ 400 K on a timescale of several picoseconds^{30–33}. Thereafter, further cooling occurs through slow coupling to acoustic phonons on timescales of ~ 100 ps to 10 ns (depending on carrier density)^{30,31}.

The second mechanism, which is not relevant to pump–probe measurements of isolated graphene sheets without contacts^{30–33}, is

energy transport away from the junction by the carriers. First, the hot carriers drift out of the junction under the built-in electric field. This occurs on a timescale of ~ 40 – 400 ps in our devices (Supplementary Fig. S9). Second, the Peltier effect cools the junction because of the large difference in Seebeck coefficient across the junction. Numerical simulations show that the PC decays on a timescale of ~ 5 ps through this mechanism (Supplementary Fig. S9). The Peltier cooling rate is proportional to the photo-thermoelectric current, which is inversely proportional to the RC time constant of the circuit. This suggests that τ should increase as the device resistance increases: indeed, the two appear to be correlated, as shown in Fig. 3b. Hence, the evidence points to Peltier cooling playing an important role in hot-carrier relaxation at the p–n junction.

Unlike coupling to optical phonons, which cools the whole laser excitation area, the Peltier mechanism only cools the junction region. The two mechanisms may therefore be distinguished in future studies using high-resolution techniques. We note that Peltier cooling has already been found to play a significant role in the cooling of graphene–metal contacts using scanning thermal microscopy³⁴.

We can use the above transient carrier transport picture to understand the dramatically different temperature dependences of the PC amplitude under the pulsed and c.w. excitations apparent in Fig. 3c. Such a temperature dependence under c.w. excitation

