Joule-heating induced thermal voltages in graphene three-terminal nanojunctions

Pascal Butti,1,a) Rolf Brönnimann,1 Klaus Ensslin,2 and Ivan Shorubalko1,b)

1) Empa - Transport at Nanoscale Interfaces, Swiss Federal Laboratories for Materials Science and Technology, 8600 Dübendorf, Switzerland
2) Solid State Physics Laboratory, ETH Zurich, 8093 Zurich, Switzerland

(Dated: 14 March 2018)

Intrinsic voltage rectification is investigated in a graphene three-terminal nanojunction (GTTJ) on Si/SiO2 at room temperature and 87 K. Room-temperature rectification efficiency (ratio of output against input voltage) reaches \( \approx 40\% \), which is higher than most efficiencies reported in the literature. The observed efficiency is higher at room temperature than at 87 K, which is in contrast to field-effect simulations and indicates that other mechanisms contribute to the rectification effect. We propose an explanation based on Joule heating and thermal voltages, as the device is operated in regimes of substantial power dissipation. Predicted thermal voltages show temperature, bias- and gate-voltage dependences which are similar as in our experiment. We conclude that Joule-heating effects need to be considered for GTTJ devices.


a)Electronic mail: pascal.butti@alumni.ethz.ch
b)Electronic mail: ivan.shorubalko@empa.ch
Nonlinear electrical properties of nanosized three-terminal junctions (TTJs) attract scientific attention because of the interest in the underlying physical phenomena and their potential for nanoelectronic applications. Early TTJs were based on III-V high-mobility semiconductor materials and operated in the ballistic or quasi-ballistic regime up to room temperature\textsuperscript{1-11}, which is highly promising for high-frequency (THz) signal processing\textsuperscript{12-15}. A number of compact nanoelectronic devices such as frequency multipliers\textsuperscript{16}, logic gates\textsuperscript{17,18}, set-reset latches\textsuperscript{19}, and half-/full-adders\textsuperscript{20,21} were demonstrated based on the functionality of these TTJs. However, the complexity of III-V technology complicates industrial production of these devices. Silicon is a more suitable material for this purpose but it suffers from comparably low mobility\textsuperscript{22}. Graphene on the other hand is not only technologically CMOS compatible but can also show high mobility at room temperature\textsuperscript{23,24}. After the first observation of intrinsic voltage rectification in graphene-based TTJ by Jacobsen et al.\textsuperscript{25}, a dozen of studies on such devices with (mainly) different types of graphene and different device geometries have been reported\textsuperscript{26-35}. Importantly, all aforementioned GTTJ devices operate in the diffusive charge-carrier transport regime. While the rectification functionality of GTTJs is robust, reported efficiencies show large variation (~0.5 - 40%). A diffusive field-effect transistor model of GTTJs predicts a few % rectification efficiency at room temperature and a slight increase at liquid nitrogen temperature\textsuperscript{36}. Recently, four-terminal nanojunctions based on graphene encapsulated between hexagonal boron nitride showed very high responsivity at room temperature, attributed to the ballistic nature of the device\textsuperscript{37}. Earlier, it was identified that the efficiency of ballistic rectifiers can be enhanced due to the thermopower effect\textsuperscript{38-41}. However, high rectification efficiency in diffusive GTTJs at room temperature is not well understood and no consensus on physical mechanisms has been attained.

In this study, voltage rectification is investigated at 296 K and 87 K in a GTTJ on Si/SiO\textsubscript{2}. Efficiencies up to almost 40% are demonstrated at room temperature, which is higher than most efficiencies reported in the literature. Strikingly, the efficiency is higher at room temperature than at 87 K, which is in contrast to field-effect simulations and indicates that other mechanisms contribute to the rectification effect. We propose an explanation based on Joule heating and thermal voltages, as the device is operated in regimes of substantial power dissipation. The predicted thermal voltages show temperature, bias- and gate-voltage dependences similar as observed in our experiment.

The device was made from exfoliated graphene flakes on top of highly doped oxidized
FIG. 1. (a) GTTJ device and measurement configuration (main figure is 4x4 μm² atomic force micrograph). Graphene terminals are pale blue, SiO₂ dark blue, and electrical contacts red. Graphene edges are indicated by dotted lines. A push-pull bias ±V_{in} was applied to two terminals, generating a current I. Two SMUs with force (F) and sense (S) outputs formed a four-point measurement. V_{out} was probed with a multimeter at the remaining inner contact. The remaining outer contact was left open. A back-gate voltage V_G was applied to the silicon substrate. Inset in the lower left corner shows a 500x500 nm² zoom-in of the central part of the device. The width of the constrictions is ≈100 nm (red scale), etched depth ≈1.5 nm. (b,c) Voltage-rectification functionality at 296 K and 87 K. V_{out} as function of V_{in} for gate voltages corresponding to hole transport (up-bending curves), Dirac point (flat curves), and electron transport (down-bending curves).

silicon chips (285 nm SiO₂). Raman spectroscopy was used to identify mono-layer graphene flakes. The flake was scanned with atomic force microscopy to check cleanliness and flatness. Metal contacts and TTJ geometry were made with standard lithography techniques. Graphene devices fabricated by such a technique typically have a few thousand cm²/√s charge carrier mobility at room temperature. The geometry consisted of three triangular terminals [Fig.1(a)], joined by 100 nm wide constrictions. Measurements were made in a N₂ environment at atmospheric pressure. The device had two contacts per terminal: one to
apply voltage and another one closer to the junction to probe voltage. In this way, contact resistances which may depend on voltage bias were excluded from the measurements. Contact resistances have so far not been excluded in publications on GTTJs, with the exception of Jacobsen et al. The electrical characterization equipment consisted of source measurement units (SMUs) Keithley 236 and a Keithley 2000 multimeter. The SMUs have force and sense outputs. The applied (force) voltage is adapted such that the probed (sense) voltage is equal to the voltage that the user wants to apply. Simultaneously, current is measured by an amperemeter. In order to avoid the influence of hysteretic effects on the studied rectification effect, the gate voltage was swept at a rate of 1-3 V/sec at which hysteresis was suppressed. At 87 K no hysteresis was observed.

The voltage-rectification functionality of the GTTJ is presented in Fig.1(b,c): under a push-pull bias \( \pm V_{in} \) on two terminals, the voltage on the remaining terminal (output voltage \( V_{out} \)) is positive when operating in the hole regime (positive rectification), negative when operating in the electron regime (negative rectification), and zero at the Dirac point (no rectification). These dependencies remain unchanged for opposite sign of \( V_{in} \). The type of rectification functionality can be adjusted by gate-voltage (Fermi-energy) tuning. \( V_{out} \) shows a quasi-parabolic dependence on \( V_{in} \) for gate-voltages away from the Dirac point (\( \approx 2 \) V). The curves are largely symmetric with respect to the sign of \( V_{in} \). Efficiency \( |V_{out}/V_{in}| \) is larger at room temperature than at 87 K (respectively \( \approx 40\% \) and 10\% for \( V_{in} = 400 \) mV).

Unwanted contributions to \( V_{out} \) stemming from device asymmetry are removed by calculating its odd and symmetric parts. It can be shown that the odd part contains contributions from asymmetry and that the symmetric part (\( V_{out,\text{sym}} \)) contains the rectification effect. Asymmetry can be caused for example by geometry or inhomogeneous doping. Fig.2 shows the gate-voltage dependence of \( V_{out,\text{sym}} \) for different input voltages. \( V_{out,\text{sym}} \) swings smoothly from positive (hole transport) to negative (electron transport). In both transport regimes, \( |V_{out,\text{sym}}| \) generally has a maximum value at a gate voltage close to the Dirac point and decreases for gate voltages further away from the Dirac point.

Field-effect simulations cannot satisfactorily explain the experimental results [Fig.3(c)]. At room temperature, a maximum output voltage of \( \approx 17 \) mV was measured for \( V_{in} = 100 \) mV, whereas for simulations it is an order of magnitude smaller. The experimental results at 87 K can be partly understood by the simulation results. These findings indicate that the rectification observed in our experiments is caused by other mechanisms.
The conductance\textsuperscript{43} $G = |I|/|2V_{in}|$ shows two clear trends for increasing input voltage [Fig.3(a)]: an increase of minimum conductance and a decreasing slope at gate voltages away from the Dirac point (i.e., decrease in carrier mobility). Applying electric fields of the order of $V/\mu\text{m}$ to graphene leads to drift-velocity saturation\textsuperscript{41-46}, which can explain the mobility decrease. Moreover, large current densities (up to 0.5 mA/\mu m) are reached in the constrictions, which is in the range of breakdown\textsuperscript{47,48}. Joule heating of the constrictions can therefore be expected. We attribute the conductance increase around the Dirac point to a Joule-heating induced increase of thermally excited charge carriers\textsuperscript{49}.

Several groups have investigated the relationship between temperature rise and power density (dissipated electric power per graphene surface area) in graphene field-effect transistors on Si/SiO\textsubscript{2} at room temperature\textsuperscript{50-54}. Joule heat in graphene on SiO\textsubscript{2} is widely thought to be dissipated vertically through the bulk SiO\textsubscript{2}\textsuperscript{44,46,50-54}. Reported data sets show a large spread in temperature. The device from Chae \textit{et al.}\textsuperscript{51} is most similar to our device in terms of channel sizes and by the fact that it was measured in a four-point configuration. Therefore, we compare calculated temperature rises (see below) to data from this device [Fig.3(b)].

The dissipated power density is largest at the constrictions, leading to the formation of a hot spot. The heat diffusion length along graphene\textsuperscript{51,55} is of the order of 100 nm and can thus be neglected. Furthermore, radiation heat loss is negligible. The hot spot area over which power was dissipated is $\approx 0.1 \mu\text{m}^2$. A power density of $\approx 0.02 \text{mW}/\mu\text{m}^2$ was reached at $V_{in} = 100 \text{mV}$, which gives already an appreciable temperature rise of $\approx 100 \text{K}$.

A rough temperature estimate is obtained from the conductance data. For this, the
conductance at the Dirac point, $G_{\text{Dirac}}$ is assumed to be proportional to the total carrier density at the Dirac point: $G_{\text{Dirac}}(T) \propto 2 n_{th}(T) + n^*$, where $T$ is the temperature, $n_{th}(T) = \frac{\pi}{6} \left( \frac{k_B T}{\hbar v_F} \right)^2$ the thermal carrier density\textsuperscript{49}, and $n^*$ the residual carrier density due to potential disorder. Furthermore, conductance is assumed to be proportional to carrier mobility $\mu(T)$ (equal for electrons and holes). The conductance curve for lowest input voltage (negligible Joule heating) was associated with the chip temperature $T_0$. By rearranging the ratio $G_{\text{Dirac}}(T)/G_{\text{Dirac}}(T_0)$, the temperature $T$ is obtained

$$T = \frac{\hbar v_F}{k_B} \sqrt{\frac{3}{2 \pi} \frac{G_{\text{Dirac}}(T)\mu(T)}{G_{\text{Dirac}}(T_0)\mu(T_0)} (2n_{th}(T_0) + n^*) - n^*}. \quad \mu(T) \text{ and } \mu(T_0) \text{ are obtained from linear fits to the conductance curves for gate voltages away from the Dirac point, taking the average of electron and hole mobility.}$

$n^* \approx 10^{11} \text{ cm}^{-2}$ is estimated from a log-log-plot of conductance vs. carrier density\textsuperscript{56}. Calculated temperatures reach as high as 400 K above the bath temperature, plotted in Fig.3(b), it agrees well with temperatures reported by Chae et al.\textsuperscript{51}. Temperature rise is comparable at 296 K and 87 K, pointing to similar Joule-heating mechanisms at both temperatures.

In the following, our experimental findings are analyzed in a picture incorporating Joule heating and thermal voltages. Due to high currents the constrictions heat up significantly (temperature $T$) above chip temperature $T_0$. The electrical contacts are thermal anchors for the graphene system, such that the temperature at the contacts is assumed equal to $T_0$. Due to the temperature difference between the TTJ center and the contact used to probe the output voltage, a thermal voltage $V_{th}$ builds up between those regions. The rectification effect due to the field effect is neglected and thus the voltage in the TTJ center is assumed to be zero. Thus, under this assumption the measured output voltage is $V_{th}$. In spite of a temperature difference between the TTJ center and the biased terminals, no thermal voltage builds up in those terminals because the voltages are fixed (four-point measurement configuration).

$V_{th}$ is calculated from the Seebeck coefficient\textsuperscript{57} $S$. In linear-response approximation $S = -dV/dT$. For metals at low temperatures ($k_B T \ll E_F$), $S$ is given in the Boltzmann picture by the Mott-formula: $S_{\text{Mott}} = -\frac{\pi^2 k_B^2 T}{3e} \left. \frac{G'(E)}{G(E)} \right|_{E=E_F}$, where $k_B$ is the Boltzmann constant, $e$ the fundamental charge, $T$ temperature, $G$ conductance, $E$ energy, and $E_F$ the Fermi energy. This relation was used successfully to explain the experimentally determined $S$ in graphene\textsuperscript{58,59}. The Mott-formula remains largely accurate up to room temperature\textsuperscript{57-59}. We use it to calculate an estimate of $S$ at $V_{in} = 100 \text{ mV}$ (roughly linear regime and $T$ roughly
FIG. 3. (a) Conductance as function of gate voltage for different $V_{in}$-values (in steps of 100 mV) at 87 K. $e^2/h$ is the conductance quantum. (b) Temperature vs power density at 296 K and 87 K, compared to Chae et al. adapted with permission from51. Copyright (2010) American Chemical Society. (c) Comparison of output voltages for $V_{in} = 100$ mV at 296 K and 87 K: experimental data ($V_{out,sym}$), calculated thermal voltages ($V_{th}$), and field-effect simulations36. All curves are shifted to a Dirac point at 0 V. In both cases, the temperature rise is roughly 90 K.

296 K or 87 K, respectively). For simplicity, $S_{Mott}$ is calculated assuming constant temperature. Maximum values are $\approx 150$ µV/K at room temperature and $\approx 50$ µV/K at 87 K, roughly in agreement with literature values. Using $S \propto T$ (according to the Mott formula), $V_{th}$ is obtained by $V_{th}(T) = -\int_{T_0}^{T} S(T')dT' = \frac{S(T_0)}{2T_0} (T^2 - T_0^2)$, where $T$ is the temperature in the TTJ center and $T_0$ the chip temperature. $T$ is determined from Fig.3(b) at a power density for a maximum value of $|V_{out,sym}|$. Fig.3(c) shows a comparison of experimental output voltages and $V_{th}$. Good agreement is seen at room temperature, and also at 87 K.
Thermal voltages therefore offer an explanation at relatively low bias. Rectification efficiency tends to saturate with increasing bias, which can be understood by Seebeck-coefficient saturation. According to theoretical literature, the graphene Seebeck coefficient saturates at high temperatures \( (k_B T > E_F) \). For typical \( E_F \approx 50 \text{ meV} \), this corresponds to \( T \sim 600 \text{ K} \). The (maximum) saturation value is \( S_{sat} \approx 200 \text{ µV/K} \). For room temperature data at \( V_{in} = 400 \text{ mV} \), the dissipated power at maximum efficiency is \( \approx 30 \text{ µW} \), giving a power density \( \approx 0.3 \text{ mW/µm}^2 \). According to Fig.3(b), \( T \) is then \( \approx 700 \text{ K} \).

Finally, the thermal voltage \( V_{th,max} = S_{sat} (T - T_0) \approx 80 \text{ mV} \). This value is only half the maximum \( V_{out,sym} \). These results could imply that \( S_{sat} \) is larger than 200 \text{ µV/K}, or that the temperatures in the constrictions were actually higher than estimated from conductance curves. Although the saturation tendency of the rectification efficiency can be understood by Seebeck-coefficient saturation, the calculated thermal voltages do not fully explain the output voltages at high \( V_{in} \).

In conclusion, voltage rectification was experimentally demonstrated at 296 K and 87 K in a GTTJ of 100 nm constriction size. While at 87 K rectification efficiency can be partially explained by the field effect, at room temperature a large discrepancy between field-effect simulations and experimental data is observed. Thus, other mechanisms seem to contribute to the rectification. Joule-heating induced thermal voltages were identified here as a possible contributing mechanism. Such thermal mechanisms must be relevant, especially in high-bias regimes where high power dissipation densities \( (\sim \text{ mW/µm}^2) \) are reached. Joule heating offers a possible explanation for high efficiency at room temperature and comparatively low efficiency at 87 K. Temperature-rise estimates (several hundred K) are in agreement with values reported in the literature, and the corresponding thermal-voltage estimates can partially explain the output voltages at room temperature and 87 K. We conclude that Joule-heating effects need to be considered for GTTJ devices. This work also suggests that multi-terminal nanojunctions can serve as an experimental playground to investigate thermo-electric effects in graphene at high temperatures, in constrictions, and at large thermal gradients.

The authors thank FIRST-lab at ETH Zurich for access to nanofabrication facilities, Empa and ETH for access to their atomic force microscope and Raman setups. Financial support by the Swiss National Science Foundation (SNF) Project Nr. 200020_159204 and
REFERENCES

17. H. Q. Xu, I. Shorubalko, D. Wallin, I. Maximov, P. Omling, L. Samuelson, and W. Seifert,


For simplicity, the linear conductance is used. We found that an analysis with differential conductance (determined numerically) leads to very similar results.


