Graphene nanoribbon field-effect transistors on wafer-scale epitaxial graphene on SiC substrates

Wan Sik Hwang,1,2,† Pei Zhao,1 Kristof Tahy,1 Luke O. Nyakiti,3,4 Virginia D. Wheeler,3 Rachael L. Myers-Ward,3 Charles R. Eddy, Jr.,3 D. Kurt Gaskill,3 Joshua A. Robinson,5 Wilfried Haensch,6 Huili (Grace) Xing,1 Alan Seabaugh,1 and Debdeep Jena1,2
1 Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA
2 Department of Materials Engineering, Korea Aerospace University, Goyang City, Gyeonggi-do 412791, South Korea
3 U. S. Naval Research Laboratory, Washington, DC 20375, USA
4 Department of Materials Science and Engineering, Texas A&M University, College Station, Texas 77843, USA
5 Materials Science and Engineering & Center of 2D & Layered Materials, Pennsylvania State University, University Park, Pennsylvania 16802, USA
6 IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA
(Received 10 July 2014; accepted 15 December 2014; published online 5 January 2015)

We report the realization of top-gated graphene nanoribbon field effect transistors (GNRFETs) of ~10 nm width on large-area epitaxial graphene exhibiting the opening of a band gap of ~0.14 eV. Contrary to prior observations of disordered transport and severe edge-roughness effects of graphene nanoribbons (GNRs), the experimental results presented here clearly show that the transport mechanism in carefully fabricated GNRFETs is conventional band-transport at room temperature and inter-band tunneling at low temperature. The entire space of temperature, size, and geometry dependent transport properties and electrostatics of the GNRFETs are explained by a conventional thermionic emission and tunneling current model. Our combined experimental and modeling work proves that carefully fabricated narrow GNRs behave as conventional semiconductors and remain potential candidates for electronic switching devices. © 2015 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [http://dx.doi.org/10.1063/1.4905155]

The implementation of 2-dimensional (2D) graphene for digital logic devices has proven challenging because of the material's zero band gap.1 Various alternate digital logic device structures have been proposed which take advantage of interlayer tunneling, graphene-3D semiconductor hetero-structures, and properties that exploit the light-like energy dispersion of carriers in 2D graphene.2–7 From the point of view of realizing conventional field-effect transistors, well-controlled graphene nanoribbons (GNRs) mimic the excellent electrostatic properties of carbon nanotubes (CNTs) and offer hope for graphene-based digital logic devices.8 The ultrathin body enables scaling down to 10 nm or below while still keeping short-channel degradation effects at bay. GNRs suffer from edge-roughness scattering effects compared with those of CNTs, but GNRs provide better large-area scalability, planar fabrication opportunity, and heat dissipation capacity.9 The availability of broken bonds at the edges provides a window of opportunity for chemical doping,10 which remains difficult in CNTs due to saturated sp2 chemical bonds. A number of “beyond-CMOS” devices, such as the GNR tunneling field-effect transistor (TFET),11 can be realized if controlled GNRs are fabricated on large-area substrates. Thus, progress in the fabrication and

†Invited for the Two-dimensional Materials special topic.
Electronic addresses: whwang@kau.ac.kr and djena@nd.edu
characterization of wafer-scale GNRs stands to potentially enable a host of applications in the future.

The creation of controlled band gaps by quantum confinement of carriers in GNRs remains a significant challenge.\textsuperscript{12-21} To date, graphene nanoribbon field effect transistors (GNRFETs) ranging down to 10–20 nm channel width have been fabricated from exfoliated graphene\textsuperscript{13,14} and chemical vapor deposition (CVD) grown graphene\textsuperscript{15,16} using conventional top-down lithography and etching methods. Bottom-up techniques such as chemically derived GNRFETs down to sub-5 nm width have been fabricated and show substantial band gaps with $I_{\text{ON}}/I_{\text{OFF}} \sim 10^6$ at room temperature.\textsuperscript{17} GNRFETs have also been fabricated by unzipping CNTs.\textsuperscript{18-20} More recently, GNRs down to 5 nm has been directly grown on SiC substrates using ion implantation followed by laser annealing.\textsuperscript{21} But the bottom-up techniques are not yet site-controlled or reproducible, and are currently incompatible with conventional lithographic processes for circuit implementations.

Epitaxial graphene (EG) grown on single-crystal, semi-insulating SiC wafers satisfy many of the above criteria.\textsuperscript{22} Furthermore, devices based on EG require fewer processing steps and are more immune to contamination compared to CVD-grown large-area graphene due to the absence of a transfer process.\textsuperscript{23} GNRFETs can mimic properties of carbon nanotube field-effect transistors (CNTFETs) and remove needs of alignment and randomization of metallic and semiconducting channels. The major challenge in realizing GNRs is in achieving ~5 nm widths with smooth edges. In this pursuit, GNRFETs stand to benefit from recent process developments in Silicon FinFET technology in which arrays of ~5 nm wide Si fins have been demonstrated with robust structural integrity.\textsuperscript{24} Process variation challenges of such narrow fins have been addressed for next-generation of CMOS technology.\textsuperscript{25}

Despite the importance of EG, substantial energy gaps have not yet been demonstrated in GNRFETs made in EG on SiC.\textsuperscript{26} Furthermore, there are no studies that correlate experimentally measured transport properties and theoretical models for EG-GNRs. In this work, we report the fabrication of top-gated ~10 nm wide GNRFETs by lithography on large area EG on SiC substrates. We observe for the first time, the opening of a substantial energy gap inversely proportional to the GNRFET width of EG-GNRs. By relating the measured transport with theoretical modeling, we find that the transport properties of narrow epi-GNRs are similar to well-behaved narrow-bandgap semiconductors, contrary to carrier localization effects reported extensively in wider GNRs fabricated on exfoliated graphene.\textsuperscript{27-31} The reasons for these observations will be discussed.

The starting material in this work was epitaxial-graphene grown on a 4 in. diameter Si-face 6H-SiC substrate. The epitaxial growth conditions are described in earlier reports\textsuperscript{22} and this

![FIG. 1.](image)
epitaxial graphene on SiC is expected to have a lower residual charge than transferred graphene (2–5 x 10^{11} cm^{-2}) on SiO_2 due to the absence of a transfer process.\cite{ref3, ref4} Figure 1 shows the final device images including a single GNR and arrays of GNRs. HSQ, a negative-tone electron-beam resist, was used to fabricate GNRs of varying widths, down to ~10 nm. The gate stack consists of 15 nm HSQ followed by atomic layer deposited (ALD) 30 nm Al_2O_3 at 200 °C and Cr (5 nm)/Au (100 nm) using an electron-beam evaporator. The GNRs are connected to the two-dimensional (2D) graphene area, in order for the energy barrier for either electrons or holes entering the GNR to be half of the GNR band gap and symmetry. The source/drain contact metal of Cr (5 nm)/Au (100 nm) sits on the 2D area forming an ohmic contact with a zero band gap. From the transfer characteristics, extrinsic field-effect mobility was extracted at 800–1000 cm^2/V s at maximum transconductance. The contact resistance was not accounted for in the mobility extraction and the contact resistance extracted from transmission line method (TLM) patterns was around 10^3 Ω μm. Details of the HSQ process and device processing flow have been discussed earlier.\cite{ref5}

Figure 2(a) shows the measured drain current I_D as a function of the gate bias V_{GS} of 10 nm GNRFETs at three different temperatures. The gate modulation (ratio of I_{ON}/I_{OFF}) of the drain current is about 10×. The relatively high I_{OFF} observed at 300 K is due to a thermionic emission current from the source contact. For a 10 nm wide GNR, the energy gap is E_g ~ 0.14 eV, which leads to a Schottky barrier height of qφ_B ~ E_g/2 ~ 70 meV. This is only slightly smaller than ~ 3 kT at room temperature, implying a large thermionic emission current over the barrier where \( I_{em} \sim \exp[-qφ_B/kT] \). This temperature dependence is accentuated at lower temperatures, because \( I_{em} \) stays relatively constant whereas \( I_{eff} \) is reduced by several orders of magnitude due to the reduction of the thermionic emission current. This results in an increase of \( I_{em}/I_{off} \rightarrow 10^6 \) at 4 K as shown in Fig. 2(a). At this low temperature, the Fermi-Dirac tail of the electron distribution in the source is severely curtailed and electrons have to tunnel through the energy gap of the GNR. This band-to-band transport mainly happens across the barrier formed at the contact, but not across the entire length of the device.

This strong temperature dependence of I_D in the GNR FET shown in Fig. 2(a) is distinctly different from that of the 2D FETs,\cite{ref6} revealing the presence of an energy gap. The family I_D versus V_{DS} of 10 nm GNRFETs in Fig. 2(b) clearly shows the “turn-on” and “turn-off” region depending on the location of the Fermi level, which is tuned by the gate bias, V_{GS}. Figure 2(c) shows I_D vs. V_{GS} for GNRs of three different widths (10, 13, and 17 nm) at a V_{DS} biased near the charge neutral voltage. The current-voltage curve in Fig. 2(c) is a characteristic of back-to-back metal-semiconductor Schottky diodes, and the turn-on window is a measure of the Schottky barrier height. It is observed that as the widths of GNRs decrease the size of the low-conductance window increases. The energy gap is inversely proportional to the widths of GNR FET channels. In order to measure the band gap quantitatively, a more comprehensive approach entails measuring the conductance map as a function of V_{DS} and V_{GS}. The results of such measurements are discussed next.
FIG. 3. The differential conductance of two representative GNRFETs of 10 nm width (a) and 17 nm width (b) as a function of $V_{DS}$ and $V_{GS}$ at 4 K. Modeling results of two different widths of 10 nm (c) and 17 nm (d) GNRFETs which are corresponding to (a) and (b), respectively. The black (dark) color represents a low conductance as indicated by the color map. (e) The energy band diagram was used in the model which was developed based on the Schottky barrier. (f) The extracted band gap of GNRFET vs. width of GNR. The linear line was predicted using the model. The deviation of these GNRs width is around 0.5 nm by SEM.

Figure 3 shows the 4 K conductance versus $V_{DS}$ and $V_{GS}$ for two representative GNRFETs of 10 nm (a) and 17 nm (b). The conductance is shown as a color in a logarithmic scale with red representing high (on-state) and black as low conductance (off-state). For a fixed drain bias, scanning the gate voltage results in a sharp transition from conducting to insulating states. For example, the 4 K scan in Fig. 3(a) is for a drain bias of 20 mV in which the transition is seen in the region $-8 \, \text{V} < V_{GS} < -6 \, \text{V}$. Similarly, for a fixed gate bias, scanning the drain voltage reveals the back-to-back Schottky behavior as shown in Figs. 3(a) and 3(b). Realizing that this channel is no different from a traditional semiconductor (albeit with a small band gap), an accurate method to extract the energy band gap is to model the entire conductance map using traditional semiconductor transport equations, accounting for both thermionic emission and tunneling current components. Because the contacts are Schottky barriers of height half of the energy band gap, and tunneling depends on the band gap, modeling the dependence of the conductance maps on the GNR widths enables an accurate extraction of the energy band gap.

The details of the hybrid thermionic emission/tunneling transport model are provided in the supplementary material accompanying this paper. The modeled conductance versus $V_{DS}/V_{GS}$ for GNRs of widths 10 nm and 17 nm is shown in Figs. 3(c) and 3(d) alongside the measured experimental results. The simple textbook-model of transport captures the entire shape of the conductance map. Note that no localization or quantum-dot type hopping transport was used in this model. The band-edge fluctuations that may result from the line-edge roughness of the GNRs cause a smearing of the on-off state transition that is observed in the experimental data as compared to the sharp transitions predicted by the model. Such fluctuations are a measure of the disorder in the GNR, but these are minimal compared to the overall characteristics, which are captured in the band-transport picture. We note that such fluctuations are not limited to GNRs alone; indeed, such fluctuations exist on any material system due to the residual charge distribution and impurities. However, their effect lessens for wide gap semiconductors while this fluctuation becomes pronounced and noticeable in narrow bandgap semiconductors. We do not observe any Coulomb diamonds or charging effects as reported earlier. The reason for this difference from earlier reports is twofold. First, because the GNRs reported here are among the narrowest reported to-date fabricated by lithographic techniques, the corresponding energy gaps are among the highest. Second, the epitaxial graphene on SiC substrates reveals smaller potential variation of 12 meV than that of transferred graphene on SiO$_2$. 
of 59–77 meV\textsuperscript{27–33} The potential fluctuations due to charged impurities can localize carriers if the band gap is small, whereas a larger band gap coupled with low residual impurity density enables conventional band-edge transport.

The dependence of the band gaps on the GNR widths extracted from Figs. 3(a) and 3(b) is shown in Fig. 3(f) and is in good agreement with a conventional model of GNR band gaps.\textsuperscript{40} An energy band gap of \( E_g \approx 0.14 \) eV was achieved by scaling to a GNR width of ~10 nm. Based on this observation, an \( E_g \approx 0.3 \) eV could be achieved by narrowing the GNR width down to 5 nm, a potential possibility in the future using FinFET technology.\textsuperscript{24} Since the transport model is based on thermionic emission and tunneling, it may be used for predicting the behavior of GNRFETs of different channel lengths. A crucial test for a transport model is tunneling from the source to drain being heavily dependent on the S/D separation at small drain voltages, whereas thermionic emission is not dependent. Using the model and the corresponding experimental measurement, we can verify the accuracy of the model further. To do so, we performed transport measurements on GNRFETs with varying S/D distances and compared them with the predictions from the model; which will be discussed next.

Figures 4(a) and 4(b) show the \( I_D \) vs \( V_{GS} \) of 15 nm wide GNRFETs for three gate lengths: 5 \( \mu \)m, 1 \( \mu \)m, and 0.1 \( \mu \)m measured at 300 K (a) and 4 K (b). It is observed that \( I_D \) increases as the gate length decreases at both 300 K and 4 K. As the gate length decreases, the gate modulation remains relatively constant at 300 K in Fig. 4(a), whereas the gate modulation changes exponentially at 4 K in Fig. 4(b) since the conduction is dominated by tunneling. The corresponding predictions based on the hybrid thermionic emission/tunneling current model are shown as solid lines in Figs. 4(a) and 4(b). The energy band diagram corresponding to the model is shown in Fig. 4(c). The model captures most of the experimentally observed behavior, further lending credence to the claim that transport in the GNRs is band-like, and both hopping and localization effects do not need to be invoked in order to explain the device behavior. As shown in the supplementary material, \( T_S \) and \( T_D \) are two coefficients determined by the source and drain barriers and \( V_{GNR} \) represents the local potential of the GNR channel.

Several recent studies have associated charge transport in GNRs with hopping conductivity and quantum dot behavior,\textsuperscript{27–31} and not by conventional conduction mechanisms. We discuss these earlier findings in light of our observations. The observations can be resolved by paying careful attention to GNR widths, surface potential variation of graphene, GNR edge roughness, and device operation regimes. First, in earlier reports, GNR widths ranging from 30 to 100 nm, which will lead to energy band gaps less than ~50 meV. This energy gap is comparable to the electron-hole puddle surface potential variations, which have been reported to be around 50–80 meV\textsuperscript{27–31} for the
graphene/SiO\textsubscript{2} interface, but only 12 meV for graphene/SiC interface.\textsuperscript{39} When the disorder potential variations are of the order of or more than the energy band gap, it constitutes a severe perturbation of transport properties. Furthermore, the high density of such fluctuations in graphene/SiO\textsubscript{2} interfaces exacerbates the localization of carriers leading to hopping transport. On the contrary, when the energy gap is larger, as obtained with narrower GNRs, the potential disorder behaves like a weak fluctuation, similar to ionized impurity doping in traditional semiconductors. The residual charge densities of GNRFETs in previous reports\textsuperscript{27–31} are expected to be high, since the GNRs were fabricated from exfoliated graphene transferred on to SiO\textsubscript{2} substrates. In addition, the HSQ mask (15 nm height) produced by EBL to etch graphene results in smooth epi-graphene GNRs, whose edge roughness is estimated to be less than \(\sim 0.35\) nm using root mean square (RMS) estimation of the width by image processing.\textsuperscript{15} Edge roughness from previous work is around 4 nm.\textsuperscript{28} Finally, the device operation regime in the conductance map reported in this work spans hundreds of meV range, unlike \(\sim 50\) meV ranges reported earlier. The experimental results reported here and the above discussions suggest that even though there is a potential fluctuation caused by either line edge roughness or potential inhomogeneity, the behavior of epi-GNR FETs is indeed no different from any conventional narrow-bandgap semiconductor. Most effects such as the ratio of \(I_{\text{ON}}/I_{\text{OFF}}\) observed in the transport of previously reported GNRs mimic those of disordered or heavily doped narrow-bandgap semiconductors,\textsuperscript{41} and as GNRs become narrower and cleaner, their intrinsic properties and electrostatic advantages will make them highly attractive for electronic devices in the future.

The array GNRFETs consisting of parallel arrays of 30 GNRs were fabricated and the \(I_D\) versus \(V_{GS}\) of the 30-array GNRFET was compared with that of a single GNRFET, all with the same 13 nm width as shown in Figs. 5(a) and 5(b). The inset of Fig. 5(b) shows the schematic image of an array GNRFET, comparing the inset of a single GNRFET in Fig. 5(a). It shows that the individual device performance of array GNRFET is preserved in the array structure. The increase of the drain current is one of the benefits of array GNRFETs. We also observe a high maximum drain current density of \(\sim 12\) mA/\(\mu\)m considering the total channel width. Such high current drives have never been reported in any semiconductor device.\textsuperscript{42} If we consider only the active ribbon width, the maximum high drain current density becomes 28 mA/\(\mu\)m, a value that may be approached by changing the pitch of the GNR array. We attribute this high current carrying capability to the high electrical and thermal conductivity of the GNR channels due to the absence of lateral scattering,
coupled with the excellent thermal conductivity of the underlying SiC substrate. The high current drives are attractive from many viewpoints: for high-performance transistors with fast switching and possibly for integrated interconnects.

Finally, it is also worth mentioning the crystallographic direction of GNRs. It is predicted that GNRs with armchair shaped edges (AGNR) can be either metallic or semiconducting depending on their widths, and that GNRs with zigzag shaped edges (ZGNR) are metallic with peculiar edge states on both sides of the ribbon regardless of their widths. Indeed, the origin of the bandgap in nanoribbons is still under debate. However, experimental observations show that band gaps are inversely proportional to the GNRs’ width with no sign of crystallographic direction dependence.

In this work, we also did not observe any signature of band gap dependence on the crystallographic direction. We believe that there are several reasons why this dependence is not observed experimentally; first, channel direction in the transistor is neither AGNR nor with atomically precise control, so it is safe to mention that GNRs possess mixed properties of both AGNR and ZGNR. Second, even though we have perfectly aligned AGNR and ZGNR in the entire channel, the GNRs' properties will be highly dominated by the edge termination structure rather than the overall crystallographic direction. In this work, the edge of epi-GNRs is fully hydrogen terminated since the GNRs are covered by HSQ, which possesses plenty of hydrogen. Based on these understandings at this present moment, the energy gap is created due to the charge carrier quantization and therefore we treat GNRs like well-behaved narrow-bandgap semiconductors.

In summary, we report results of the first top-gated 10 nm width GNRFETs on a large-area epitaxial graphene exhibiting exceptionally high drive currents, the opening of a substantial band gap and an increase of drain current by exploiting FET arrays. The narrow GNR width in the range of 10 nm and the epitaxial platform enables a conventional current flow mechanism without introducing the hopping effect or quantum dot behavior. The measured transport dependence over the entire parameter space (GNR width, gate length, temperature) is explained accurately by invoking a single conventional thermionic emission + tunneling model. With further scaling of the widths of wafer-scale clean GNRFETs, graphene based transistors can show promising potential for practical applications.

This work was supported by the Semiconductor Research Corporation (SRC), Nanoelectronics Research Initiative (NRI), and the National Institute of Standards and Technology (NIST) through the Midwest Institute for Nanoelectronics Discovery (MIND). STARnet, an SRC program sponsored by MARCO and DARPA, and by the Office of Naval Research (ONR), and the National Science Foundation (NSF). L.O.N. is grateful for the postdoctoral support from the ASEE and W. S. Hwang acknowledges the support from Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2014R1A1A004770).

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://aplmaterials.aip.org/about/rights_and_permissions

35. See supplementary material at http://dx.doi.org/10.1063/1.4905155 for the details of the hybrid thermionic emission/tunneling transport model.