SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
ELECTRONIC PROPERTIES AND DEVICES
APPLIED PHYSICS
SCALING LAWS
NANOSENSORS

Received
2 January 2013

Accepted
10 January 2013

Published
19 February 2013

Correspondence and requests for materials should be addressed to J.L. (jinglu@pku.edu.cn)

Sub-10 nm Gate Length Graphene Transistors: Operating at Terahertz Frequencies with Current Saturation

Jiaxin Zheng[1,2,3], Lu Wang[3], Ruge Quhe[1,2], Qihang Liu[1], Hong Li[1], Dapeng Yu[1], Wai-Ning Mei[3], Junjie Shi[1], Zhengxiang Gao[1] & Jing Lu[1]

[1]State Key Laboratory of Mesoscopic Physics and Department of Physics, Peking University, Beijing 100871, P. R. China, [2]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, P. R. China, [3]Department of Physics, University of Nebraska at Omaha, Omaha, Nebraska 68182-0266.

Radio-frequency application of graphene transistors is attracting much recent attention due to the high carrier mobility of graphene. The measured intrinsic cut-off frequency (f_T) of graphene transistor generally increases with the reduced gate length (L_{gate}) till L_{gate} = 40 nm, and the maximum measured f_T has reached 300 GHz. Using *ab initio* quantum transport simulation, we reveal for the first time that f_T of a graphene transistor still increases with the reduced L_{gate} when L_{gate} scales down to a few nm and reaches astonishing a few tens of THz. We observe a clear drain current saturation when a band gap is opened in graphene, with the maximum intrinsic voltage gain increased by a factor of 20. Our simulation strongly suggests it is possible to design a graphene transistor with an extraordinary high f_T and drain current saturation by continuously shortening L_{gate} and opening a band gap.

Graphene is of particular interest for ultrahigh speed electronics due to its high carrier mobility and saturation velocity[1–6]. Because of its zero band gap, graphene field effect transistors (FETs) have a low on/off current ratio, which limits its application in logic devices. There is still a lack of reliable techniques to open a sizable gap without degrading the electronic properties of graphene. However, a large on/off ratio is not necessary for radio frequency (r. f.) electronic applications[6], which are the core elements in wireless communication devices. Developing the high-performance r. f. graphene transistors is attracting enormous recent attention[7–16]. In order to facilitate high-performance r. f. applications, FETs should respond quickly to the gate voltage (V_g), which requires short gates[6]. Extensive gate length (L_{gate}) scaling work of graphene FETs has been experimentally pursued[9–14], with the gate length scaling down to 40 nm[14]. The intrinsic cut-off frequency (f_T) represents how fast the channel current is modulated by the gate and is one of the most important figure-of-merit for evaluating the performance of r. f. devices. All these experimental investigations reveal that f_T of graphene transistors generally increases with the decreasing L_{gate}. The measured maximum f_T is 300 GHz for a graphene FET with L_{gate} = 144 nm based on exfoliated graphene[10]. The maximum f_T estimated from the static measurement is 1.4 THz from a graphene FET with L_{gate} = 45 nm fabricated by a self-aligned approach[9]. Therefore, as far as f_T is concerned, graphene FETs have significantly outperformed conventional silicon metal-oxide-semiconductor field effect transistors (MOSFETs) with a highest measured f_T of 485 GHz at L_{gate} = 29 nm[6,17] and III-V high-electron-mobility transistors (HEMTs) with a highest measured f_T of 660 GHz at L_{gate} = 20 nm for GaAs[6].

However, the f_T value does not always increase with the reduced L_{gate} in a transistor when L_{gate} approaches the size limit. For example, f_T of GaAs metamorphic HEMT peaks at L_{gate} = 20 nm and then it slightly decreases with the reduced L_{gate}[6]. One fundamental issue arises naturally: is there a saturation of f_T with the reduced L_{gate} in graphene FETs? If such a saturation is absent, a higher intrinsic cut-off frequency, even up to tens of terahertz (THz), can be obtained via continuously shortening L_{gate} to a few nm in graphene FETs. Besides f_T, the maximum oscillation frequency (f_{max}) and intrinsic voltage gain (A_v) are the other two important figure-of-merits evaluating the r. f. performance of devices. To obtain a high f_{max} and A_v, a drain current saturation is required. Although a drain current saturation has been reported in some monolayer graphene (MLG) FET devices due to phonon scatter limited velocity saturation[18] and in a dual-gated bilayer graphene (BLG) FET device due to an electrical-field-induced band gap opening[19], all the existing short-channel graphene FETs with L_{gate} below 300 nm suffer

from a lack of drain current saturation[12,14], which lead to a small output conductance and a poor f_{max} and A_v. Another fundamental issue of graphene FETs is: is there an effective method to induce a drain current saturation in sub-10 nm scale?

In this article, we investigate theoretically for the first time the performance limit of top-gated graphene FETs with the gate length scaling down from 9.86 to 0.91 nm using the density functional theory (DFT) coupled with the nonequilibrium Green's function (NEGF) method. We demonstrate that switching effects remain in such short gate graphene transistors. Remarkably, f_T still increases with the decreasing L_{gate}, with values of 3.4–21 THz. The absence of a saturation in f_T with the reduced L_{gate} clearly shows that f_T of graphene transistors can be continuously improved by shortening L_{gate}. In order to create a drain current saturation, we design two schemes: one is a dual-gated FET made of a MLG sandwiched between two hexagonal boron nitride (h-BN) layers and the other is a dual-gated FET made of a pure BLG. A significant current saturation has been observed in both FETs under a proper vertical electrical field, and the maximum voltage gain has been increased up to 2.3 and 1.5, respectively, a factor of 20 and 13 higher than that of pure MLG FETs. Experimentally, sub-10 nm gate length carbon nanotube transistors have been recently manufactured via top-down approach[20]. It has been pointed out by Duan *et al.* that a sub-10 nm gate length graphene FET can be realized by using the self-aligned approach with a sub-10 nm nanowire as the top gate[9]. Finally, we envision an alternative method to fabricate a sub-10 nm graphene transistor, namely, using a sub-10 nm diameter boron nitride (BN) or carbon nanotube as the top gate. We expect our work can inspire experimentalists to further push graphene transistors to performance limit by continuously shortening L_{gate} and opening a band gap.

Results

In an actual graphene FET, metal is always used to contact graphene as electrodes. In our schematic model of a MLG FET presented at Fig. 1, we design the graphene channel contacted underneath two aluminum (Al) electrodes (source and drain). The distance between Al contact

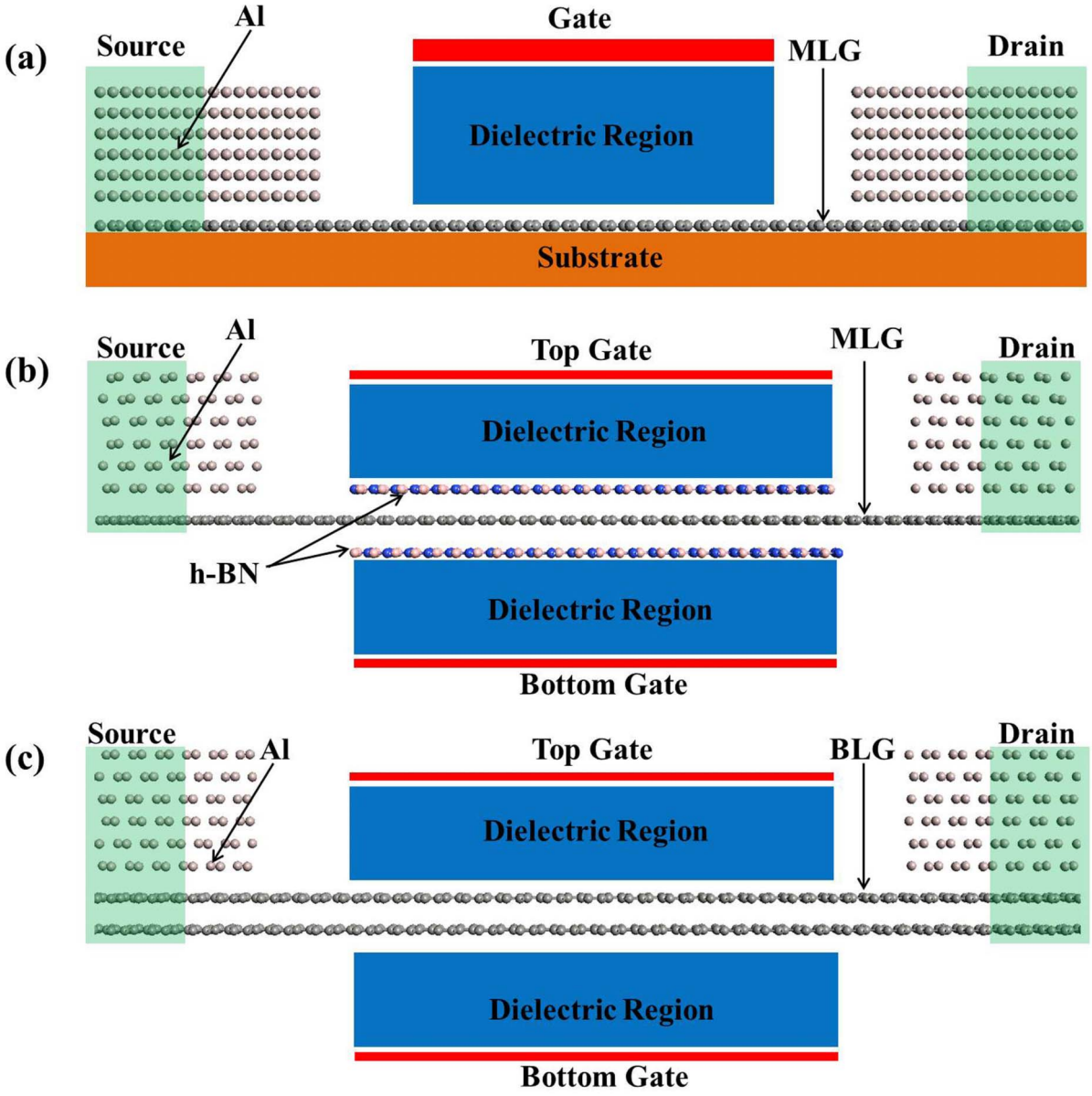

Figure 1 | (a-c) Schematic model of a top-gated pure MLG FET (a), dual-gated BN/MLG/BN sandwich FET (b), and dual-gated BLG FET (c). The two-dimensional MLG and BLG channels are contacted underneath two aluminum (Al) electrodes. Gray ball: C; light pink ball: Al; blue ball: B; heavy pink ball: B.

and graphene layer is 0.34 nm, according to the previous work[21,22]. We further set the thickness of the dielectric region to be $d_i = 1.4$ nm, and the dielectric constant to be $\varepsilon_r = 3.9$, which models after SiO_2.

As the gate of FETs scales down to very short lengths, like sub-10 nm region, the transport mechanism turns from channel-dominated diffusive regime to the contact-dominated ballistic regime, so the contact effect between the metal electrode and the channel becomes very important for the device transport properties[20,23]. We first study the contact effect between the aluminum (Al) electrode and graphene. In Fig. 2(a) and (b), we show the projected density of states (PDOS) on the carbon (C) atoms of graphene in the left/right lead ($g_{L/R}(E)$) and in the channel ($g_{ch}(E)$), respectively, for a graphene FET with $L_{gate} = 5.6$ nm under zero drain-source voltage and zero gate voltage. The inset in Fig. 2(a) is the density of states (DOS) of the pure MLG graphene, with the Dirac point exactly at the Fermi level (E_f). We find that the Dirac points of graphene in the left/right lead ($D_{L/R}$) and in the channel (D_{ch}) are shifted to $D_{L/R} = E_f - 0.9$ eV and $D_{ch} = E_f - 0.25$ eV, respectively, indicative of heavy n-doping of graphene in the contact part and light n-doping in the channel part by Al contact. This contact effect is also reflected on the transmission spectrum of this device. As shown in Fig. 2(c), there are two transmission minima in the transmission spectrum of this graphene FET but only one minimum in the transmission spectrum of the corresponding MLG FET without being contacted underneath Al electrodes (See the inset in Fig. 2(c)). The transmission coefficient of the device $T(E)$ behaves like three resistors composed of the channel and the two electrodes in series[24–26]:

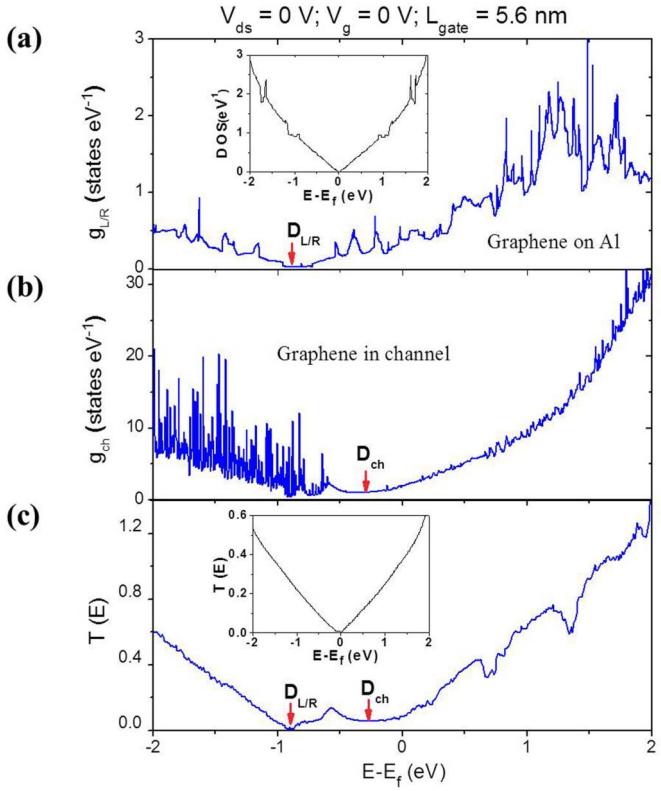

Figure 2 | Contact effect between the graphene channel and metal electrodes for a 5.6-nm gate length MLG FET under zero drain-source voltage and zero gate voltage. (a, b) Projected density of states on carbon atoms of graphene in the left/right lead (a) and in the channel (b). (c) Transmission spectrum of this device. $D_{L/R}$ and D_{ch} denote the Dirac point of the graphene in the left/right electrode and the channel graphene, respectively. Inset in (a): density of states for the pure MLG; inset in (c): transmission spectrum of MLG without being contacted underneath metal electrodes under zero drain-source voltage and zero gate voltage.

$$T(E) \propto \frac{g_{ch}(E)g_L(E)g_R(E)}{g_{ch}(E)g_L(E)+g_{ch}(E)g_R(E)+g_L(E)g_R(E)} \quad (1)$$

The two minima in the transmission spectrum originate from the $D_{L/R}$ in $g_{L/R}(E)$ and D_{ch} in $g_{ch}(E)$, respectively. Consequently, there is a lack of symmetry in the transmission spectrum with respect to both $D_{L/R}$ and D_{ch}, causing the asymmetry in electron and hole conductance. The existence of two minima in the transmission spectrum and electron-hole conductance asymmetry agrees well with both previous theoretical and experimental works[24,25,27–30].

Fig. 3(a) shows the transmission spectra of this device under $V_g = 6.0, -3.0$, and -5.0 V with a fixed bias voltage of $V_{ds} = 0.5$ V. There are three transmission minima (D_L, D_R, and D_{ch}) in each transmission spectrum because D_L and D_R are separated by the bias voltage. D_L and D_R are unchanged by V_g. However, D_{ch} is shifted by V_g: first it is at E_f under $V_g = -5$ V, then shifted to $E_f - 0.25$ eV under $V_g = -3$ V, and finally shifted to $E_f - 0.9$ eV under $V_g = 6$ V. This shift of D_{ch} with V_g in the transmission spectrum stems from the shift of D_{ch} with V_g in $g_{ch}(E)$, as shown in Fig. 3(b). As D_{ch} is at E_f ($V_g = -5$ V), the $T(E)$ values within the bias window are small, and the current is minimized (off-state). As D_{ch} is moved gradually out from the bias window, the $T(E)$ values within the bias window increase, the currents increase gradually, and the FET is finally switched from the off-state ($V_g = -5$ V) to the on-state ($V_g = 6$ V). The on/off current ratio of this device is 13.3.

We calculate the transmission eigenvalues of all the k-points in the 1D Brillouin zone at E_f. For a given energy E, the transmission coefficient $T(E)$ is given by

$$T(E) = \int_{-\pi/a}^{\pi/a} \frac{dk}{2\pi} \sum_n T_{nn}(E,k) = \int_{-\pi/a}^{\pi/a} \frac{dk}{2\pi} \sum_n \lambda_n(E,k) \quad (2)$$

where λ_n are the eigenvalues of the transmission matrix T. In Fig. S1 of Supplementary Information, we show the transmission coefficients $T(E_f, k)$ as a function of k of the off-state ($V_g = -5.0$ V, black curve) and the on-state ($V_g = 6.0$ V, red curve) in the 1D Brillouin zone. The width and height of the $T(E_f, k)$ spectrum for the on-state are much larger than those for the off-state, and therefore the total transmission factor $T(E_f)$ of the on-state is larger than that of the off-state. To visualize the difference between the off- and on-state of this graphene FET, we choose the $(2\pi/5a, 0)$ point of the k-space at E_f and show its transmission eigenchannel in the off-state (left panel) and on-state (right panel) in Fig. 3(c). The largest transmission eigenvalue at this point of the off-state is $\lambda_1 = 1.1 \times 10^{-3}$, and the rest are nearly zero; correspondingly, the incoming wave function is almost completely scattered and unable to reach the other lead. In contrast, the largest transmission eigenvalue at this point of the on-state is $\lambda_1 = 0.97$, and the rest are also nearly zero; consequently, the incoming wave function is scattered little and most of the incoming wave reaches the other lead.

We show the transfer characteristics of the devices with L_{gate} scaling down from 9.86 to 0.91 nm in Fig. 4(a), at a fixed drain-source voltage of $V_{ds} = 0.5$ V. As expected, all of them show an n-type feature. The on-currents, which are defined as the current at $V_g = 6.0$ V, remain almost constant for all the L_{gate}. By contrast, the leakage currents in the off-state increase with the decreasing L_{gate}. As shown in Fig. 4(b), the on/off current ratio drops significantly from 15.8 at $L_{gate} = 9.86$ nm to 3.35 at $L_{gate} = 0.91$ nm, and this scaling trend agrees with the experimental results for short gate graphene FETs ($L_{gate} = 5.6$ um – 50 nm)[12,23]. Such a scaling behavior is attributed to the increasing off-state leakage current with the decreased L_{gate}: as the gate scales down to very short lengths, the off-state average PDOS on each C atom of the channel graphene at D_{ch} increases (Fig. S2, Supplementary Information), and the channel graphene behaves more like metal and away from semi-metal. Our calculated on/off ratio of MLG FETs is larger than the previous

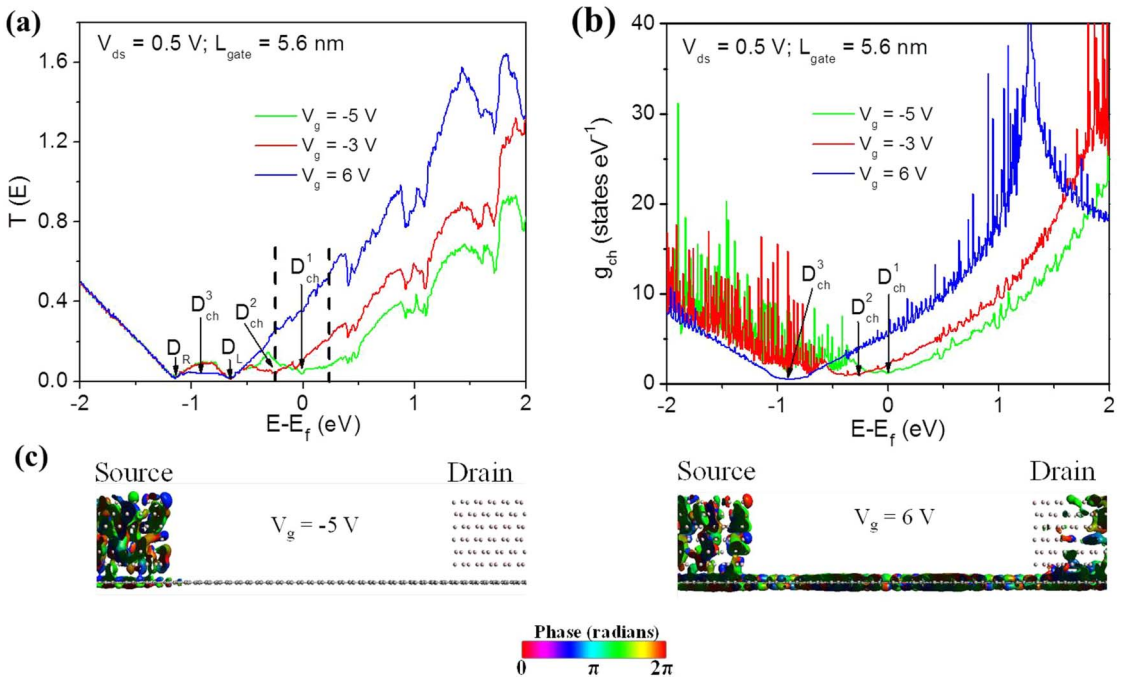

Figure 3 | Switching effect for a 5.6-nm gate length MLG FET with a drain-source voltage of V_{ds} = 0.5 V. (a) Transmission spectra of the off-state (V_g = −5.0 V), on-state (V_g = 6.0 V), and an intermediate state (V_g = −3.0 V).The dashed vertical line indicates the bias window. $D_{L/R}$ denotes the Dirac point of the graphene in the left/right electrode; D^1_{ch}, D^2_{ch}, and D^3_{ch} denote the Dirac point of the channel graphene under V_g = −5.0, −3.0, and 6.0 V, respectively. (b) Projected density of states on carbon atoms in the channel under V_g = −5.0, −3.0, and 6.0 V, respectively. (c) Transmission eigenstates of the off-state (V_g = −5.0 V) and on-state (V_g = 6.0 V) at E_f and k = (0.4, 0). The isovalue is 0.6 a.u.

experimental data (For example, in the work of Wu *et al.*, the on/off current ratios are about 3.5–6 for p-type graphene FETs with L_{gate} = 50–500 nm)[23]. This difference is mainly ascribed to the fact that the thickness of the dielectric region in our model is only 1.4 nm, which is much thinner than those (90 nm in Ref. 30) of the experimental devices. The effect of the thickness of the dielectric region on the performance of graphene FETs has been studied in detail by Guo *et al.*[31]. They find that a thinner dielectric region will improve the on/off current ratio and transconductance significantly because the gate modulation is more effective as the dielectric region becomes thinner. The method to fabricate an ultrathin dielectric (sub-10 nm) will be discussed in the next section.

Transconductance g_m is another important parameter to characterize switching effect of an electronic device and can be extracted from the transfer characteristics. In Fig. 4(c), we present its dependence on L_{gate}. The g_m values first decrease slowly from 3670 μS/μm at L_{gate} = 9.86 nm to 2900 μS/μm at L_{gate} = 2.19 nm and then start to drop dramatically to 1296 μS/μm at L_{gate} = 0.91 nm. The reduced g_m with the reduced L_{gate} agrees with both previous experimental and theoretical results[13,31] and is not favorable for f_T. We attribute the reduced g_m to the increasing leakage current and the shift of Dirac voltage (V_{Dirac}, defined as the gate voltage at the point of the off-state). The increasing leakage current is apparent. The shift of V_{Dirac} can be seen in the transfer characteristics shown in Fig. 4(a). The shift is remarkable when L_{gate} scales below 2.19 nm (from V_{Dirac} = −8 to −22 V as L_{gate} scales from 2.19 to 0.9 nm), which is responsible for the dramatic drop of g_m. The main reason for the shift of V_{Dirac} is the short channel effect[32,33]: when the L_{gate} is very small, the electrostatic potential profile is strongly affected by the drain-source potential, and a bigger gate voltage is needed to turn off the channel. We notice that g_m = 3670 μS/μm of our results at L_{gate} = 9.86 nm is larger than 2300 μS/μm for a self-aligned graphene FET with L_{gate} = 90 nm at V_{ds} = 1.0 V reported experimentally by Duan *et al.*[9]. This difference is also owing to the extremely thin dielectric region in our model compared with the experiment.

To get f_T, we need to calculate another key parameter: the intrinsic gate capacitance C_g. As shown in Fig. 4(d), C_g decreases approximately linearly with L_{gate}, because C_g is related to L_{gate} through the following equation[14]: $C_g = \varepsilon_0\varepsilon_r W_{gate}L_{gate}/t_{ox}$, where ε_0 is the dielectric constant of vacuum, ε_r the relative dielectric constant of the gate dielectric, W_{gate} the gate width, and t_{ox} the gate dielectric thickness. Actually, besides C_g, transistors also include another kind of capacitance named parasitic capacitance (C_p), which exists between the gate and source (drain) electrode and is an important parameter to determine the extrinsic cut-off frequency ($f_{T,\,ex}$): $f_{T,\,ex} = g_m/(C_g + C_p)$[15]. Ideally, one can make $f_{T,\,ex}$ close to f_T through reducing the parasitic capacitance, such as by redesigning the device layout and reducing the total parasitic series resistance between the device's source (drain) electrode and the gate to increase the current density[10,34,35].

Fig. 4(e) shows L_{gate} dependence of f_T based on the calculated g_m and C_g. f_T increases monotonically from 3.4 to 21 THz as L_{gate} scales down from 9.86 to 0.91 nm, because C_g reduces faster than g_m with the reduced L_{gate}. We notice that f_T exhibits a $1/L_{gate}$ dependence as L_{gate} scales from 9.86 to 2.19 nm, which agrees well with the scaling trend reported experimentally by Wu *et al.* with L_{gate} from 550 to 40 nm[14]. The $1/L_{gate}$ scaling trend suggests that the transport of our devices is in the contact-dominated ballistic regime and the electric field along the channel is dominated by the value of the contact resistance at the device's source (drain)[14]. This $1/L_{gate}$ dependence is also usually observed for short-channel conventional Si and III-V FETs[14]. As L_{gate} scales below 2.19 nm, the scaling trend deviates from the $1/L_{gate}$ dependence, due to the dramatic drop of g_m. The product $f_T L_{gate} = g_m t_{ox}/2\pi\varepsilon_0\varepsilon_r W_{gate}$ is expected to be linearly proportional to g_m for devices with the same parameters and dimensions. In Fig. 4(f), we show the plot of $f_T L_{gate}$ against g_m, and it exhibits the expected linear dependence.

We summarize the gate length scaling works of the intrinsic cut-off frequency for MLG FETs in Fig. 5. An increase trend of f_T with the reduced L_{gate} is available in all works. Almost all the data can be

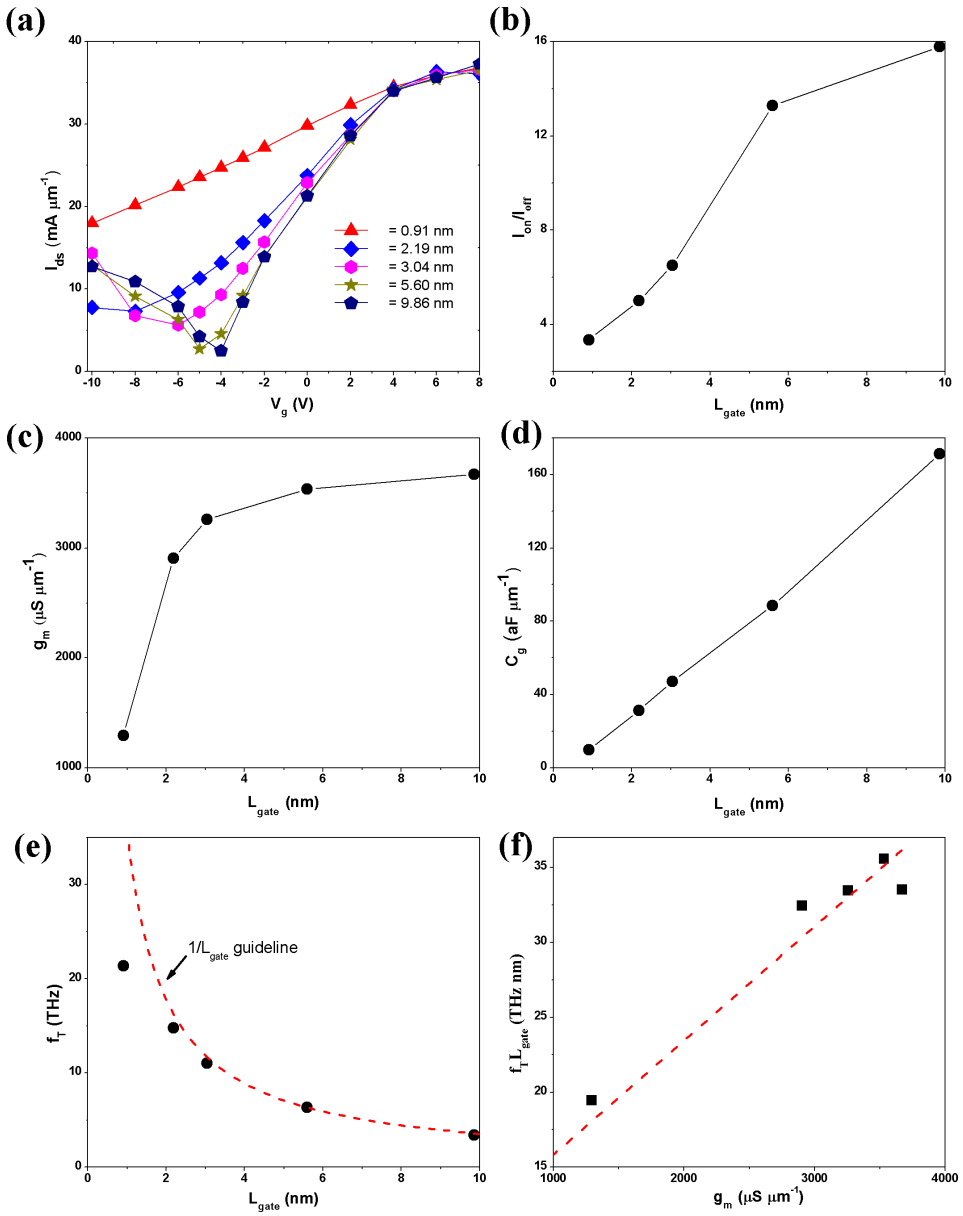

Figure 4 | (a-e) Gate length scaling of the MLG FETs: transfer characteristics (a), on/off current ratio (b), transconductance obtained from the transfer characteristics at $V_g = 0$ V (c), intrinsic gate capacitance (d), and intrinsic cut-off frequency at $V_g = 0$ V (e). Gate dielectric thickness is $t_{ox} = 1.4$ nm, dielectric constant $\varepsilon_r = 3.9$, and drain-source voltage $V_{ds} = 0.5$ V. The red dashed line in (e) shows a $1/L_{gate}$ dependence. (f) Product of the intrinsic cut-off frequency and the gate length as a function of the transconductance. The red dashed line is a linear fitting.

roughly fitted by the uniform $f_T = a/L_{gate}$ relation ($a = 38600$ GHz nm) except that the data based on CVD grown graphene has a different coefficient ($a = 8000$ GHz nm). It is noteworthy that in the 40–100 nm region, the theoretical f_T data obtained by Guo *et al.* using a self-consistent ballistic quantum transport simulation approach with the NEGF formalism[31] agrees well with those reported by Duan *et al.* with a self-aligned approach[9]. This agreement shows the reliability of the quantum transport simulation in predicting f_T of graphene FETs, and the condition will be discussed in the next section.

We study the output characteristics of a graphene FET with $L_{gate} = 6.4$ nm and show them in Fig. 6(a). The source-drain ballistic current increases linearly with the applied bias voltage in the checked bias region. In most experimental graphene FET devices, the absence of drain current saturation degrades the maximum oscillation frequency (f_{max}) and the intrinsic voltage gain (A_v), which represent how fast power transmission is modulated by the gate and the amplification factor of an input signal, respectively. f_{max} is defined as the frequency at which the power gain becomes unity, and a typical approximation for it is $f_{max} = \frac{f_T}{2\sqrt{g_d(R_g + R_{ds}) + 2\pi f_T R_g C_g}}$[36], where g_d is the output conductance, R_g the gate resistance and R_{ds} the drain-source resistance. A_v is defined as $A_v = g_m/g_d$.

The reason of the absence of current saturation in graphene FET devices is attributed to zero band gap of graphene, high electrical resistance between the device's source and drain electrode and its gate, and the short channel effect[6,32,33,37]. One way to get current saturation is to open a band gap in graphene[6]. Both theoretical and experimental works show that a vertical external electric field can induce a tunable band gap up to 0.25 eV for BLG without degrading the electronic properties of graphene[38–40], and a current saturation in BLG FETs ($L_{gate} = 4 \sim 9$ μm) is indeed observed and reproduced in a simulation ($L_{gate} = 40$ nm) within the tight-binding Hamiltonian

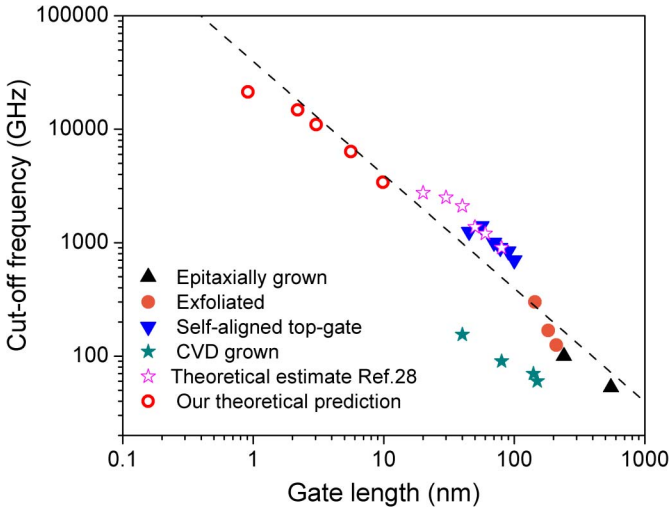

Figure 5 | Gate length scaling of the intrinsic cut-off frequency for different graphene FETs. The experimental values are based on the epitaxially grown graphene[11], exfoliated graphene[10], self-aligned nanowire gate[9], and CVD grown graphene[14], respectively. Data are roughly fitted by the curve showing a $1/L_{gate}$ dependence.

and NEGF formalism as a result of the band gap opening[19,40]. Our previous theoretical work predicts that a vertical external electric field can induce a tunable band gap up to 0.34 eV for a MLG properly sandwiched between two h-BN single layers without degrading the electronic properties of graphene[41]. Such a BN/MLG/BN sandwich structure has been prepared experimentally recently[42].

The schematic model of a dual-gated graphene BN/MLG/BN sandwich FET with L_{gate} = 6.4 nm and a dual gated BLG FET with L_{gate} = 9.6 nm are presented in Fig. 1 (b) and (c), and the two individual gates allow us to create a larger band gap and tune channel's conductance individually. The device performance of MLG and BLG FET with the same L_{gate} is similar if the band gap of BLG is not opened[19]. The vertical electrical field applied to sandwich structure is obtained as. The corresponding total gate voltage is $V_g = V_t + V_b$, reflecting the total doping level. In Fig. 6(b), we compare the simulated output characteristics of a pure MLG FET and a BN/MLG/BN sandwich FET (under a vertical external electric field of −1 V/Å) with the same L_{gate}. A significant current saturation appears for the sandwich FET at V_g = −1.6, −2.0, and −2.4 V (a positive gate voltage has no such effect), and its output conductance g_d is lowered at most by a factor of 80 compared with the pure MLG FET. The calculated C_g of this sandwich FET is nearly intact and g_m is a quarter of those of the pure MLG FET (its transfer characteristics are shown in Fig. S3(a)), and the f_T of this sandwich FET is 1.5 THz (it is degraded by a factor of 4 compared with a value of 6.35 THz for the pure MLG FET with the same L_{gate}). The maximum A_v of this

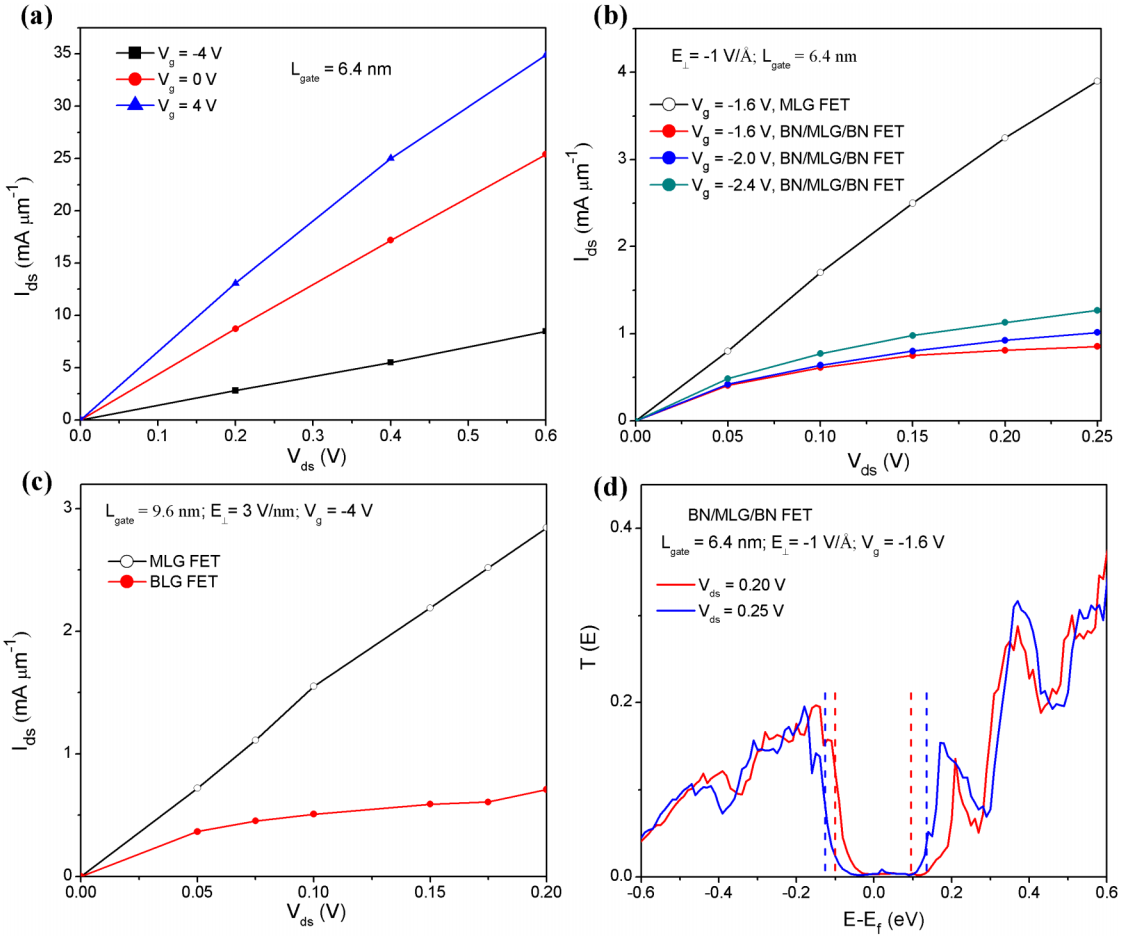

Figure 6 | (a) I_{ds}-V_{ds} output characteristics for the top-gated MLG FETs at variable gate voltages for L_{gate} = 6.4 nm. (b) I_{ds}-V_{ds} output characteristics for a pure MLG FET and a BN/MLG/BN sandwich FET with the same L_{gate} = 6.4 nm under a vertical electrical field of $E_\perp$ = −1 V/Å. (c) I_{ds}-V_{ds} output characteristics for a pure MLG FET and a BLG FET with the same L_{gate} = 9.6 nm under a vertical electrical field of $E_\perp$ = 3 V/nm. (d) Transmission spectra of the BN/MLG/BN sandwich FET with L_{gate} = 6.4 nm at V_{ds} = 0.2 V (red curve) and 0.25 V (blue curve) under a vertical external electric field of -1 V/Å and a gate voltage of −1.6 V. The dashed vertical lines indicate the bias window.

device is therefore increased to 2.3, a factor of 20 higher than that of the pure MLG FET (the experimental enhancement factor and simulation enhancement factor of A_v for BLG FETs (L_{gate} = 4–9 μm experimentally and 40 nm theoretically) by a vertical external electric field are 6 and 10, respectively)[19]. The f_{max} is increased by a factor of 2 if we assume that W_{gate} is 1 μm, R_{ds} 980 Ω, and R_g 47 Ω[16]. We also observe a significant current saturation in the output characteristics of the BLG FET at $V_g = -4.0$ V under a vertical external electric field of 3 V/nm (Fig. 6(c)), and its output conductance g_d is lowered at most by a factor of 40 compared with the pure MLG FET with the same L_{gate}. The calculated C_g of this BLG FET is nearly intact with the pure MLG FET with the same L_{gate}, and g_m is one third of those of the pure MLG FET (its transfer characteristics are shown in Fig. S3(b)), the f_T is 1.1 THz (it is degraded by a factor of 3 compared with a value of 3.4 THz for the pure MLG with the same L_{gate}), the maximum A_v is increased to 1.5, and the f_{max} is increased by a factor of 1.6. Therefore it is possible to design a sub-10 nm graphene FET that can operate at extraordinary high f_T and with greatly improved A_v by introducing a band gap in graphene. The gain factor in g_d and A_v is one order of magnitude larger than the degradation one in f_T. Besides the sandwich and bilayer scheme, adsorption of Li on MLG and formation of LiC_6 structure[43] is able to open a larger band gap of about 0.4 eV, and thus LiC_6 monolayer is also a proper candidate of the channel of MLG FET with a current saturation. Since MLG grown on insulating SiC substrate already has a band gap of 0.26 eV[44], we can alternatively use a single-gated MLG grown on SiC as the channel to create a drain current saturation although a simulation of such a device is a challenge due to the difficulty to reproduce a band gap in the calculation[45,46].

The mechanism of the current saturation in the BN/MLG/BN sandwich FET and the BLG FET can be both attributed to the band gap induced by the vertical external electric field. Taking the BN/MLG/BN sandwich FET for example, the band gap improves the electrostatic pinch-off of the channel, which is reflected on the transmission spectrum of this device (Fig. 6(d)). The center of the transport gap ($\Delta \sim 0.13$ eV) induced by the vertical external electric field is about 0.06 eV above E_f at $V_{ds} = 0.2$ V. As the V_{ds} increases from 0.2 to 0.25 V, the bias voltage shifts the transport gap to the left. Namely, the bias voltage functions like a positive gate voltage here. As a result, the integral area of the transmission coefficients within the bias window at $V_{ds} = 0.25$ V is almost the same as that at $V_{ds} = 0.2$ V, and the current is thus saturated. From the point view of transport theory, the ability of a bias voltage to shift the transmission spectrum is as important as the opening of a band gap of graphene to create a drain current saturation. If the bias voltage does not shift the transmission spectrum, the current will not be saturated. The similar drain current saturation mechanism of the BLG FET is provided in Supplementary Information (Fig. S4).

Discussion

To fabricate short-channel graphene transistors, Duan *et al.* used an ultrathin nanowire as the top gate, and the gate length depends on the diameter of the nanowire[9,10,35]. In principle, sub-10 nm gate length graphene transistors can be fabricated by this way if a sub-10 nm diameter nanowire is used as the top-gate[9]. In view of the difficulty to fabricate sub-10 nm diameter nanowires, we suggest to use a sub-10 nm diameter BN or carbon nanotube, which is experimentally accessible, as an alternative top gate. We build a graphene and (5, 0) BN nanotube junction model (Fig. 7(a)) and simulate its I-V_{bias} characteristic curve (Fig. 7(b)). The diameter of this BN nanotube is 0.4 nm, and the band gap of it is 2.15 eV. A clear rectification is observed, and there is little leakage current between the graphene channel and the BN nanotube when the bias voltage is below 1.0 V, indicating that the BN nanotube itself can function as the local gate with the interface depletion layer in the BN nanotube as a gate dielectric[9]. We can also use sub-10 nm diameter metallic carbon nanotubes as the local gate. In this case, a few h-BN layers (sub-10 nm thick) can be used as high-quality dielectrics[16].

Two factors are critical to ensure the high performance in our sub-10 nm gate length graphene transistors: ultrathin dielectric (1.4 nm) and very small access resistance due to the very small gaps (0.32 nm) between the gated graphene channel and source/drain contacts. An ultrathin dielectric implies a better gate modulation on the channel current[31]. If we use sub-10 nm diameter semiconducting nanowires, BN nanotubes, or carbon nanotubes as the top-gate to fabricate sub-10 nm graphene transistors, the dielectric region is the depletion layer, whose thickness can be controlled by doping concentration and is obviously smaller than the nanowire/nanotube diameters. Therefore, the ultra-thin dielectric (a few nm) can be realized in such fabrication schemes. If we use sub-10 nm diameter metallic nanowires or carbon nanotubes as the local gate and a few h-BN layers (sub-10 nm thick) as dielectrics, the ultra-thin dielectric can also be realized.

A substantial access resistance due to the significant gaps between source/drain and gate electrodes (a large portion of the graphene channel in the gap area is not gated) limits the achievable transconductance and has adverse impact on short channel devices[9,47,48]. Experimentally, the access resistance has been significantly reduced with a self-alignment approach[9,10,47], through which the edges of the source, drain, and gate electrodes are automatically and precisely positioned so that no overlap and significant gaps exist between them. This process improves the transconductance and the drain current density significantly[9,10,47], and a slight current saturation was even observed[35]. As a result, this self-alignment approach finally improves the r. f. performance of short-channel graphene transistors. For example, the f_T values estimated from the static measurement of graphene FETs by Duan *et al.* in the 40–100 nm gate region have

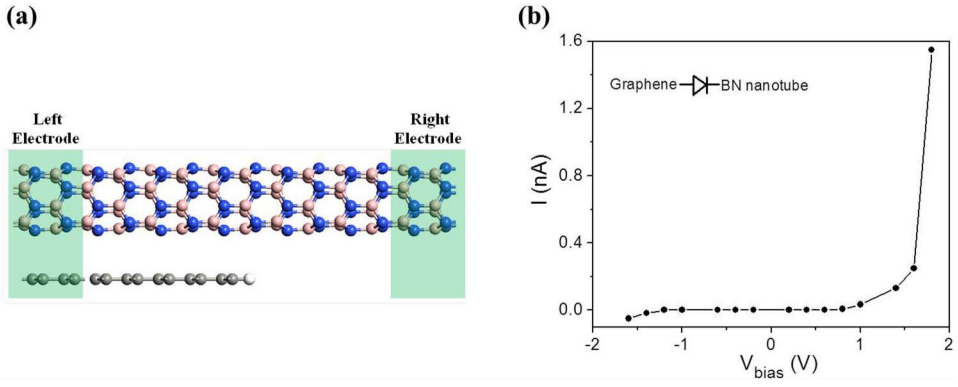

Figure 7 | Schematic model of a MLG and (5, 0) boron nitride nanotube junction (a) and its I-V_{bias} output characteristic curve (b).

reached 700 ~ 1400 GHz as a result of ultrathin dielectric and ultrasmall access resistance. Such a f_T-L_{gate} relation agrees well with the simulation reported by Guo *et al.* within a model also with an ultrathin dielectric (16 nm) and an ultrasmall access resistance due to the ultrasmall distance between the source/drain and gate electrodes[9,31].

In our model, we assume that electron transport is in the ballistic region. Another issue we need to discuss here is the probable roles of electron-phonon scattering. For carbon nanotubes, some works on this issue reveal that the measured or calculated high-field electron mean free path (MFP) is about 10 nm or longer due to optical phonon scattering and the low-field MFP 1600 nm due to acoustic phonon scattering[49–51]. The experimental and theoretical work by Dai *et al.* also points out that transport through very short (~10 nm) nanotubes is free of significant acoustic and optical phonon scattering and thus ballistic and quasiballistic at the low and high field limit, respectively[51]. For graphene, the measured electron MFP is from tens to hundreds of nanometers on rugged SiO_2 substrate, depending on the carrier concentration[52] and several micrometers in suspended ultraclean membranes[53] or in graphene encapsulated between inert and ultraflat h-BN layers at room temperature[54,55]. The channel of our graphene transistor is sub-10 nm, which is much smaller than electron MFP of graphene on smooth h-BN substrate or even on rugged SiO_2 substrate in high carrier concentration, so the transport mechanism should be safely in the ballistic regime if a smooth substrate is used, and the Landau- Büttiker transport formalism is applicable.

In summary, we have investigated the switching effect and r. f. performance of the sub-10 nm gate length graphene FETs with Al electrodes by employing *ab initio* quantum transport simulation for the first time. We find that switching effects remain in these ultrashort gate graphene transistors. The intrinsic cut-off frequency increases monotonically from 3.4 to 21 THz, which are a few to tens of times larger than the experimental maximum values of the competitive r. f. FETs[6], when the gate length scales down from 9.86 to 0.91 nm. A significant current saturation can be created in sub-10 nm graphene FETs by introducing a band gap to graphene. We expect that our theoretical work can stimulate experimental fabrication of sub-10 nm gate length graphene FETs operating at an intrinsic cut-off frequency exceeding the available best experimental values and with current saturation.

Methods

The current of graphene FETs under a finite drain-source voltage (V_{ds}) and gate voltage (V_g) is computed using the Landauer-Büttiker formula[56]

$$I(V_{ds},V_g)=\frac{2e}{h}\int_{-\infty}^{+\infty}\{T(E,V_{ds},V_g)[f_L(E-\mu_L)-f_R(E-\mu_R)]\}dE \quad (3)$$

where $f_{L/R}$ are the Fermi-Dirac distribution function for the left (L)/right (R) electrode, μ_L/μ_R are the electrochemical potential of the left (L)/right (R) electrode, and $T(E, V_{ds}, V_g)$ is the transmission coefficient. T(E) is calculated by using the Fisher and Lee relationship[57].

Then, the gate effect is calculated by solving the Poisson and Kohn-Sham equations self-consistently with the fixed boundary condition[56]. Namely, we first calculate the Hartree potential V^H from the Poisson equation: $\nabla^2 V^H = -4\pi\rho(r)$, where the initial electron density ρ(r) is computed from the DFT and NEGF methods under no gate voltage. Afterward the new Hartree potential is used to calculate a new electron density by solving the Kohn-Sham equation. These procedures are iterated until the desired numerical accuracy is reached, which at present are carried out by using the well developed ATK 11.2 package[58–60]. Single-zeta (SZ) basis set is used, the real-space mesh cutoff is 150 Ry, and the temperature is set at 300 K. The local-density-approximation (LDA) is employed for the exchange–correlation functional. The electronic structures of electrodes and central region are calculated with a Monkhorst–Pack[61] 50 × 1 × 100 and 50 × 1 × 1 k-point grid, respectively.

The intrinsic cut-off frequency f_T is defined as the frequency at which the current gain becomes unity. We calculate f_T using the flowing equation[15,31,34],

$$f_T=\frac{1}{2\pi}\frac{g_m}{C_g} \quad (4)$$

where f_T is determined by the intrinsic gate capacitance C_g and the transconductance g_m, which are computed from

$$C_g=\frac{\partial Q_{ch}}{\partial V_g}\Big|_{V_g},\quad g_m=\frac{\partial I_{ds}}{\partial V_g}\Big|_{V_g} \quad (5)$$

where Q_{ch} is the total charge of the channel, and I_{ds} the drain-source current.

In our previous work, we used the same DFT + NEGF approach to simulate the transport properties of sub-10 nm functionalized metallic single-walled carbon nanotube FETs and pentacene molecule FETs[62]. The reliability of this approach to simulate sub-10 nm FETs is verified from the fact that the order of magnitude of the calculated on/off ratio and the on-current of our results basically agree with the experimental data for the sub-10 nm single-walled carbon nanotube transistors and pentacene molecule FETs[20,63]. For example, our calculated on/off ratio and on-current for a pentacene molecule FET with L_{gate} = 0.8 nm are ~10^2 and 7.0 × 10^{-2} μA respectively, comparable with the experimental on/off ratio of ~10^3 and on-current of ~3.0 × 10^{-2} μA for a pentacene molecule FET with $L_{gate} \approx$ 1–3 nm[63].

1. Novoselov, K. S. *et al.* Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438**, 197–200 (2005).
2. Zhang, Y. B., Tan, Y. W., Stormer, H. L. & Kim, P. Experimental observation of the quantum Hall effect and Berry's phase in graphene. *Nature* **438**, 201–204 (2005).
3. Berger, C. *et al.* Electronic confinement and coherence in patterned epitaxial graphene. *Science* **312**, 1191–1196 (2006).
4. Geim, A. K. & Novoselov, K. S. The rise of graphene. *Nat. Mater.* **6**, 183–191 (2007).
5. Avouris, P. Graphene: electronic and photonic properties and devices. *Nano Lett.* **10**, 4285–4294 (2010).
6. Schwierz, F. Graphene transistors. *Nat. Nanotech.* **5**, 487–496 (2010).
7. Lin, Y. M. *et al.* Operation of graphene transistors at gigahertz frequencies. *Nano Lett.* **9**, 422–426 (2008).
8. Meric, I. *et al.* Current saturation in zero-bandgap, top-gated graphene field-effect transistors. *Nat. Nanotech.* **3**, 654–659 (2008).
9. Liao, L. *et al.* Sub-100 nm channel length graphene transistors. *Nano Lett.* **10**, 3952–3956 (2010).
10. Liao, L. *et al.* High-speed graphene transistors with a self-aligned nanowire gate. *Nature* **467**, 305–308 (2010).
11. Lin, Y. M. *et al.* 100-GHz transistors from wafer-scale epitaxial graphene. *Science* **327**, 662 (2010).
12. Bai, J. W. *et al.* Top-gated chemical vapor deposition grown graphene transistors with current saturation. *Nano Lett.* **11**, 2555–2559 (2011).
13. Meric, I. *et al.* Channel length scaling in graphene field-effect transistors studied with pulsed current-voltage measurements. *Nano Lett.* **11**, 1093–1097 (2011).
14. Wu, Y. Q. *et al.* High-frequency, scaled graphene transistors on diamond-like carbon. *Nature* **472**, 74–78 (2011).
15. Kim, K., Choi, J. Y., Kim, T., Cho, S. H. & Chung, H. J. A role for graphene in silicon-based semiconductor devices. *Nature* **479**, 338–344 (2011).
16. Meric, I. *et al.* High-frequency performance of graphene field effect transistors with saturating IV-characteristics. Presented at the Electron Devices Meeting (IEDM), IEEE, (2011) doi: 10.1109/IEDM.2011.6131472.
17. Lee, S. *et al. Tech. Dig. IEDM*, 255–258 (IEEE, 2007).
18. Barreiro, A., Lazzeri, M., Moser, J., Mauri, F. & Bachtold, A. Transport properties of graphene in the high-current limit. *Phys. Rev. Lett.* **103**, 076601 (2009).
19. Szafranek, B. N., Fiori, G., Schall, D., Neumaier, D. & Kurz, H. Current saturation and voltage gain in bilayer graphene field effect transistors. *Nano Lett.* **12**, 1324–1328 (2012).
20. Franklin, A. D. *et al.* Sub-10 nm carbon nanotube transistor. *Nano Lett.* **12**, 758–762 (2012).
21. Giovannetti, G. *et al.* Doping graphene with metal contacts. *Phys. Rev. Lett.* **101**, 026803 (2008).
22. Khomyakov, P. A. *et al.* First-principles study of the interaction and charge transfer between graphene and metals. *Phys. Rev. B* **79**, 195425 (2009).
23. Wu, Y. Q. *et al.* Quantum behavior of graphene transistors near the scaling limit. *Nano Lett.* **12**, 1417–1423 (2012).
24. Farmer, D. B. *et al.* Chemical doping and electron-hole conduction asymmetry in graphene devices. *Nano Lett.* **9**, 388–392 (2008).
25. Barraza-Lopez, S., Vanevic, M., Kindermann, M. & Chou, M. Y. Effects of metallic contacts on electron transport through graphene. *Phys. Rev. Lett.* **104**, 076807 (2010).
26. Ni, Z. Y. *et al.* Tunable bandgap in silicene and germanene. *Nano Lett.* **12**, 113–118 (2012).
27. A Zyilmaz, B. *et al.* Electronic transport and quantum hall effect in bipolar graphene p-n-p junctions. *Phys. Rev. Lett.* **99**, 166804 (2007).
28. Huard, B. *et al.* Transport measurements across a tunable potential barrier in graphene. *Phys. Rev. Lett.* **98**, 236803 (2007).
29. Williams, J. R., DiCarlo, L. & Marcus, C. M. Quantum hall effect in a gate-controlled p-n junction of graphene. *Science* **317**, 638–641 (2007).
30. Huard, B., Stander, N., Sulpizio, J. A. & Goldhaber Gordon, D. Evidence of the role of contacts on the observed electron-hole asymmetry in graphene. *Phys. Rev. B* **78**, 121402 (2008).

31. Chauhan, J. & Guo, J. Assessment of high-frequency performance limits of graphene field-effect transistors. *Nano Res.* **4**, 571–579 (2011).
32. Taur, Y. & Ning, T. H. *Fundamental of Modern VLSI Devices* (Cambridge Univ. Press, 1998).
33. Schwierz, F., Wong, H. & Liou, J. J. Nanometer CMOS (Pan Stanford, 2010).
34. Rutherglen, C., Jain, D. & Burke, P. Nanotube electronics for radiofrequency applications. *Nat. Nanotech.* **4**, 811–819 (2009).
35. Liao, L. *et al.* Scalable fabrication of self-aligned graphene transistors and circuits on glass. *Nano Lett.* **12**, 2653–2657 (2012).
36. Schwierz, F. & Liou, J. J. Modern Microwave Transistors: Theory, Disign, and Performance (Wiley-Interscience, 2003).
37. Schwierz, F. Electronics: industry-compatible graphene transistors. *Nature* **472**, 41–42 (2011).
38. Oostinga, J. B., Heersche, H. B., Liu, X. L., Morpurgo, A. F. & Vandersypen, L. M. K., Gate-induced insulating state in bilayer graphene devices. *Nat. Mater* **7**, 151–157 (2008).
39. Zhang, Y. B. *et al.* Direct observation of a widely tunable bandgap in bilayer graphene. *Nature* **459**, 820–823 (2009).
40. Xia, F. N., Farmer, D. B., Lin, Y. M. & Avouris, P. Graphene field-effect transistors with high on/off current ratio and large transport band gap at room temperature. *Nano Lett.* **10**, 715–718 (2010).
41. Quhe, R. G. *et al.* Tunable and sizable band gap of single-layer graphene sandwiched between hexagonal boron nitride. *NPG Asia Mater.* **4**, e6; doi:10.1038/am.2012.1010 (2012).
42. Britnell, L. *et al.* Field-effect tunneling transistor based on vertical graphene heterostructures. *Science* **335**, 947–950 (2012).
43. Profeta, G., Calandra, M. & Mauri, F. Phonon-mediated superconductivity in graphene by lithium deposition. *Nat. Phys.* **8**, 131–134 (2012).
44. Zhou, S. Y. *et al.* Substrate-induced bandgap opening in epitaxial graphene. *Nat. Mater* **6**, 770–775 (2007).
45. Mattausch, A. & Pankratov, O. Ab initio study of graphene on SiC. *Phys. Rev. Lett.* **99**, 076802 (2007).
46. Varchon, F. *et al.* Electronic structure of epitaxial graphene layers on SiC: Effect of the substrate. *Phys. Rev. Lett.* **99**, 4 (2007).
47. Farmer, D. B., Lin, Y. M. & Avouris, P. Graphene field-effect transistors with self-aligned gates. *Appl. Phys. Lett.* **97**, 013103 (2010).
48. Lin, Y. M. *et al.* Dual-gate graphene FETs with f_T of 50 GHz. *IEEE Electron Dev. Lett.* **31**, 68–70 (2007).
49. Yao, Z., Kane, C. L. & Dekker, C. High-field electrical transport in single-wall carbon nanotubes. *Phys. Rev. Lett.* **84**, 2941–2944 (2000).
50. Javey, A. *et al.* High-field quasiballistic transport in short carbon nanotubes. *Phys. Rev. Lett.* **92**, 106804 (2004).
51. Park, J. Y. *et al.* Electron-phonon scattering in metallic single-walled carbon nanotubes. *Nano Lett.* **4**, 517–520 (2004).
52. Giannazzo, F., Sonde, S., Nigro, R. L., Rimini, E. & Raineri, V. Mapping the density of scattering centers limiting the electron mean free path in graphene. *Nano Lett.* **11**, 4612–4618 (2011).
53. Bolotin, K. I., Sikes, K. J., Hone, J., Stormer, H. L. & Kim, P. Temperature-dependent transport in suspended graphene. *Phys. Rev. Lett.* **101**, 096802 (2008).
54. Dean, C. R. *et al.* Boron nitride substrates for high-quality graphene electronics. *Nat. Nanotech.* **5**, 722–726 (2010).
55. Mayorov, A. S. *et al.* Micrometer-scale ballistic transport in encapsulated graphene at room temperature. *Nano Lett.* **11**, 2396–2399 (2011).
56. Datta, S. *Quantum Transport: Atom to Transistor* (Cambridge Univ. Press, 2005).
57. Fisher, D. S. & Lee, P. A., Relation between conductivity and transmission matrix. *Phys. Rev. B* **23**, 6851–6854 (1981).
58. Brandbyge, M., Mozos, J.-L., Ordejon, P., Taylor, J. & Stokbro, K. Density-functional method for nonequilibrium electron transport. *Phys. Rev. B* **65**, 165401 (2002).
59. Soler, J. M. *et al.* The SIESTA method for ab initio order-N materials simulation. *J. Phys.: Condens. Matter* **14**, 2745–2779 (2002).
60. ATOMISTIX Toolkit, version 11.2. Quantum Wise A/S: Copenhagen, Denmark.
61. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).
62. Li, H. *et al.* Functionalized metallic single-walled carbon nanotubes as a high-performance single-molecule organic field effect transistor: an ab initio study. *J. Phys. Chem. C* **114**, 15816–15822 (2010).
63. Qi, P. F. *et al.* Miniature organic transistors with carbon nanotubes as quasi-one-dimensional electrodes. *J. Am. Chem. Soc.* **126**, 11774–11775 (2004).

Acknowledgements

This work was supported by the NSFC (Grant Nos. 11274016, 51072007, 91021017, 11047018, and 60890193), the National Basic Research Program of China (Nos. 2013CB932604 and 2012CB619304, MOST of China) Program for New Century Excellent Talents in University of MOE, Fundamental Research Funds for the Central Universities, National Foundation for Fostering Talents of Basic Science (No. J1030310/No. J1103205) of China, and Nebraska Research Initiative and DOE DE-EE0003174 in the United States.

Author contributions

The idea was conceived by J.L. The transport simulation was performed by J.Z. and L.W. The data analyses were performed by J.Z., L.W., R.Q., W.M. and J.L. Q.L., H.L., D.Y., J.S. and Z.G. helped discussing. This manuscript was written by J.Z., R.Q., W.M. and J.L. All authors reviewed this manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Zheng, J. *et al.* Sub-10 nm Gate Length Graphene Transistors: Operating at Terahertz Frequencies with Current Saturation. *Sci. Rep.* **3**, 1314; DOI:10.1038/srep01314 (2013).