Record Maximum Oscillation Frequency in C-face Epitaxial Graphene Transistors

Zelei Guo, Rui Dong, Partha Sarathi Chakraborty, Nelson Lourenço, James Palmer, Yike Hu, Ming Ruan, John Hankinson, Jan Kunc, John D. Cressler, Claire Berger, and Walt A. de Heer

1. School of Physics, Georgia Institute of Technology, Atlanta GA, 30332, USA
2. School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta GA, 30332, USA
3. CNRS-Institut Néel, BP166, 38042 Grenoble Cedex 9, France
*: Corresponding author

Abstract
The maximum oscillation frequency (\(f_{\text{max}}\)) quantifies the practical upper bound for useful circuit operation. We report here an \(f_{\text{max}}\) of 70 GHz in transistors using epitaxial graphene grown on the C-face of SiC. This is a significant improvement over Si-face epitaxial graphene used in the prior high frequency transistor studies, exemplifying the superior electronics potential of C-face epitaxial graphene. Careful transistor design using a high \(\kappa\) dielectric T-gate and self-aligned contacts, further contributed to the record-breaking \(f_{\text{max}}\).

Keywords
Epitaxial graphene, Silicon carbide, field effect transistors, high frequency electronics, nanoelectronics, maximum oscillation frequency

Graphene is considered a promising candidate material for high-frequency electronics. Whereas the lack of a bandgap in graphene, and consequently the small switching on-off current ratio, clearly impedes the development of graphene-based digital electronics, large on-off ratio is not necessary for some high-frequency circuits. The intrinsic low dimensionality of graphene, high carrier mobility, and large carrier velocity are clear advantages for realizing high frequency performance. From the first large scale patterning of graphene field effect transistors (GFET), performance has improved at a rapid pace. For RF transistors, the two most important small-signal figures-of-merit are the cut-off frequency \(f_t\) (the frequency at which the current gain is unity), and the maximum oscillation frequency \(f_{\text{max}}\) (the frequency at which the power gain is unity). De-embedded cut-off frequencies in graphene FETs are starting to approach the values of the best of Si-based transistors. Despite the high \(f_t\) of graphene FETs, however, transistor power gain \(f_{\text{max}}\) remains stubbornly low, limiting its capabilities for use in practical circuits. In any large-scale practical electronics application, graphene is implemented on a substrate that may greatly affect the underlying structural and electronic properties. An atomically well-defined substrate is clearly preferable over disordered substrates for most electronic applications. This has been realized by decades of transistor development on mono-crystalline silicon, where devices that are a few tens of nanometers in size are now reliably patterned by the billions/cm². Epitaxial graphene on mono-crystalline 4H- or 6H-SiC wafers has a
demonstrated large-scale patterning capability. Large wafers with atomically smooth surfaces that are ready for epitaxial growth are commercially available (at a price of about US $20/cm²). Because SiC is a wide bandgap semiconductor, epitaxial graphene does not need to be transferred to another substrate. Another advantage is that the graphene-substrate interface is formed at high temperatures, so that it is well defined, pure and reproducible.

On SiC substrates, previous efforts towards high-frequency GFET production have exclusively focused on the (0001) Si-terminated surface, although the (000-1) carbon terminated surface has reportedly higher electronic mobility. However, monolayer epitaxial graphene is more difficult to produce on the C-face than on the Si-face. Very thin films are required in this case because of the strong potential screening of the top gate by the graphene layers.

In a basic GFET structure, the average carrier density (n) in the graphene channel is modulated by the voltage applied on the gate (Vgs), which is reflected by the induced change in the current Is that flows in the channel from drain to source, and by the transconductance (gm=dIs/dVgs) (see for instance). Specifically, Is=(W/L)nev, where L and W are the channel length and width. The average carrier velocity v can be written v =1/(1/vsat + 1/µE), in which vsat is the carrier saturation velocity, µ is the carrier mobility, and E is the average electric field along the channel. For high frequencies, the current gain is defined by the ratio |H2|=|dIs/dIgs, where Igs is the gate-source current. The cutoff frequency fT is the frequency at which |H2|=1, and the maximum oscillation frequency fmax can be derived from measurements as the frequency at which the Mason’s unilateral gain U1/2=1.

For enhanced performance at high frequencies, higher carrier velocities have been pursued, which can be achieved via higher vsat or µ. For graphene there is no clear evidence for a saturation velocity per se. However, following ref. where the observed vsat is modeled by hot carriers scattering by optical phonons from the substrate (vsat = vF (hΩ/E)), with hΩ the relevant optical phonon energy), we expect a higher upper bound for vsat for epitaxial graphene on SiC (hΩ~115 meV) than for SiO₂ (hΩ~55-60 meV) similarly to graphene transferred to diamond-like carbon (hΩ~165 meV). We note that vsat could also be affected by the properties of the top-gate dielectric layer. Carrier mobilities up to µ~2000 cm²/Vs have been reported for Si-face epitaxial graphene after an hydrogenation process that results in a significant enhancement in transport properties. For top gated monolayer graphene on the C-face mobility µ~8000 cm²/Vs (Fig. 1b) is routinely achieved at high charge density, and up to 40,000 cm²/Vs at low charge density. High carrier mobility can result in lower contact resistances, which are beneficial for fmax, and would make it easier to drive the device into saturation. For example, with a mobility µ = 2000 cm²/Vs, an electric field 120 kV/cm would be required to accelerate the carriers to 75% of the saturation velocity (if limited by substrate phonons). The electric field can be reduced well below thermal limits by using high mobility C-face graphene with µ = 8000 cm²/Vs. These considerations make high quality monolayer graphene on the C-face of SiC a very good candidate for realizing high frequency FETs.

In this letter, we show that high mobility C-face GFETs built using T-gates and self-aligned contacts indeed can present promising high-frequency characteristics. In particular, their maximum oscillation frequency fmax of 70 GHz outperforms published GFETs fabricated on any other substrate materials. fT is often referred to as the most relevant figure-of-merit for
graphene RF transistors; however, \( f_{\text{max}} \) is a far more relevant parameter for most circuits, since it accounts for device parasitics and measures the frequency at which RF power can be delivered to a load.

Here, high \( f_{\text{max}} \) have been achieved by a combination of high mobility graphene and FET design. Owing to the high carrier mobility of C-face graphene, the contact resistance between graphene and metal is low (value \( R_c < 100 \ \Omega \mu \text{m} \)). The un-gated graphene resistance is minimized by utilizing a self-alignment technique, thereby minimizing the total access resistance (sum of contact resistance and un-gated graphene resistance). In addition, the self-alignment technique minimizes parasitic capacitance. T-gate structures are widely used to improve the power gain performance in III-V and silicon RF transistors, and can reduce the gate resistance, especially \(^{19}\) for short gate lengths, by incorporating a large, conductive gate head.

Monolayer C-face epitaxial graphene is grown on semi-insulating 4H-SiC single crystals using the confinement controlled sublimation method (CCS) \(^{12}\). For this approach, a 3 mm by 4 mm SiC die is inductively heated up to 1500 °C in a graphite enclosure provided with a calibrated leak. An example of a monolayer area of several hundreds of \( \mu \text{m}^2 \) growing over the SiC surface steps is shown in Figure 1. Monolayer graphene regions are identified by their Raman spectroscopy signal combined with atomic force microscopy (AFM) \(^{11,20}\). As shown in Figure 1a, the characteristic Raman spectrum of a monolayer graphene region (after subtraction of the SiC Raman peaks) comprises of a narrow G-peak (~1580 cm\(^{-1}\)) and a single Lorentzian 2D-peak (~2685 cm\(^{-1}\)). The absence of D-peak indicates that the graphene is of high structural quality. The characteristic half-integer quantum Hall effect is observed in similarly prepared monolayer graphene \(^{11,12,20}\), demonstrating the quality of the graphene monolayer on the C-face, even after deposition of a top-gate \(^{11}\). The deposition of a top gate dielectric layer (ALD Al\(_2\)O\(_3\), see below) moderately degrades the graphene mobility. High Hall and FET mobilities are consistently observed on the p- and n- side of ambipolar C-face monolayer graphene devices, as shown for instance in Figure 1b, with \( \mu_{\text{Hall}} = 7500 \ \text{cm}^2/\text{Vs} \) and \( \mu_{\text{FET}} = 8700 \ \text{cm}^2/\text{Vs} \) on the n-doped side of a 2 \( \mu \text{m} \times 4 \ \mu \text{m} \) Hall bar. The mobility \( \mu_{\text{FET}} \) was calculated using \( \mu_{\text{FET}} = (\partial \rho / \partial V_{\text{gfs}})/\rho_{\text{ox}} \), where \( \rho_{\text{ox}} \) is determined from measuring the charge induced by the gate in Hall effect measurements. These values are the highest \( \mu_{\text{FET}} \) reported so far for epitaxial graphene. Note that this graphene layer is n-doped (\( n \sim 1.6 \times 10^{12}/\text{cm}^2 \)) as expected from charge transfer from the SiC substrate \(^{1}\). However, device processing and gate deposition often compensate for the n-doping and may turn the device p-doped (as is the case for the GFET presented below in Figure 4). A schematic diagram of the GFET fabrication process flow is shown in Figure 2.

Standard lithography and metallization techniques were used, and are compatible with large-scale mass production. T-gates of various lengths were first patterned on the selected monolayer graphene regions using e-beam lithography (JEOL JBX 9300), with a tri-layer electron beam resist (Figure 2a) \(^{21}\). A thin layer of aluminum (4–5 nm) was deposited as the seed layer for subsequent low temperature atomic layer deposition (ALD) of 10 nm thick aluminum oxide as the dielectric layer (Figure 2b) \(^{22,23}\). E-beam evaporation was used to deposit 40 nm Ti / 130 nm Au on the dielectric to complete the gate stack (Figure 2c), and the remaining metal was lifted off (Figure 2d). Unlike for most wafer-scale graphene FET fabrication processes reported in the literature \(^{4,8,9}\) that involve covering the entire wafer with a dielectric layer, the above method applies the dielectric layer only to the gated region of the devices. The remaining area of graphene is left uncovered for the subsequent self-aligned
contact application. The source and drain areas were coated with 7 nm Pd/10 nm Au by angle deposition, thereby minimizing the exposed graphene area on both sides of the gate foot (Figure 2e). Residual graphene is removed with oxygen-plasma before e-beam lithographically defined source/drain/gate pads are covered with metal (40 nm Ti / 140 nm Au) (Figure 2f). For improved performance, the GFETs have a 2-finger design, with channel width 7 μm, and probing pads of 40 μm x 50 μm with a 100 μm pitch. Figure 3a shows an example of a T-gate with 100 nm foot length made on epitaxial graphene, and Figure 3b is a scanning electron microscopy image of the dual gate device.

The DC performance of the GFET devices with 100 nm and 250 nm gate lengths,respectively, are shown in Figure 4. From the output characteristics of the 100 nm gate device at low drain-source bias, we find that the unit width two-point resistance is only R_{2pt} = 200 Ωμm, indicating that the contact resistance (R_c =1/2(R_{2pt} - R_{channel})) is less than 100 Ωμm. This is the smallest per unit width 2-point resistance for graphene FETs reported to date. The contact resistance is smaller than R_c =230 Ωμm (110 Ωμm, resp.) reported for metal–graphene junction resistance at 300K (6K, resp). It is of similar value to the ultra-low resistance ohmic contacts (R_c below 100 Ωμm) reported in Ref 25 and it is comparable to that of silicon 26. Because of this small contact resistance, a large current modulation (I_{ds} vs. V_g) is observed. For instance, a current density as large as 2.6 mA/μm is observed at V_{ds}=-0.8 V and V_g=0 V. As is usual for graphene devices, the I-V characteristics of the devices doesn’t show a clear saturation, but rather an inflection point at relatively small V_{ds} 27-30. The I-V characteristics also indicate that the device is heavily p-doped, since a minimum in the I_{ds} vs V_{ds} is not reached in the 0-3 V range. In this limited V_{ds} range, large current values were measured, with I_{ds} >2.5 mA/μm at V_{ds}=-0.8 V (g_m peaks ~ 250 mS/mm), in the same range as reported in the literature for comparable bias voltages 27-29.

High-frequency scattering parameters (S-parameters) of the GFETs have been measured up to 50 GHz under ambient conditions using standard ground-signal-ground (GSG) microwave probes. The system was calibrated with the short-open-load-through (SOLT) process to eliminate the parasitic effects of the wiring and the probes 8. Before de-embedding, a maximum f_r of 41 GHz is measured for a device with 100 nm gate length at V_g=3.5 V and V_{ds}=-0.5 V (Figure 5a). This f_r value rivals the highest non de-embedded cut-off frequency reported to date for GFETs (e.g., comparable to 24). Measurements on devices of 250 nm gate length show a cut-off frequency of 33 GHz (non de-embedded), demonstrating the reproducibility of high performance of C-face graphene GFETs. To remove the effects of probing pads, we use a de-embedding procedure similar to that described in 16. Namely, the ‘open’ and ‘short’ structures have the exactly same layout as the graphene device under test (DUT), except that there is no graphene channel in the ‘open’ structure, and all the contacts are shorted to each other in the ‘short’ structure. After the de-embedding process, f_r values of 110 GHz and 60 GHz were measured on 100 nm and 250 nm gate length devices, respectively. These measured f_r values compare well with the estimated f_r = g_m/2πC_g ~ 80-110 GHz using the measured g_m =0.25mS/μm for the 100nm long and 7μm wide dual gate device, with a 15nm thick dielectric and dielectric constant κ = 6-8 for ALD-Al_2O_3. As expected, longer gate length devices have reduced f_r. Note that considering devices of similar gate length, these f_r values are higher than that of silicon MOSFETs, and about half that of HEMTs.
As discussed, a more relevant parameter for high frequency operation is $f_{\text{max}}$. We have measured $f_{\text{max}} = 38$ GHz and $f_{\text{max}} = 70$ GHz before and after de-embedding for a 100 nm gate device. To our knowledge, these $f_{\text{max}}$ values are the highest reported values to date for any GFETs, on any substrate \(^{2, 6, 30}\). High $f_{\text{max}}$ values are consistently found in our C-face graphene FETs. All six of the measured 100 nm gate length devices show $f_{\text{max}}$ after de-embedding between 50 GHz and 70 GHz. Data for two of the measured 100 nm devices are shown in Figure 5c. Together, these results show that SiC is a promising substrate for high frequency GFETs.

In prior reports on graphene RF transistors, $f_{\text{max}}$ was not only relatively small compared with $f_\tau$ of the same devices, it was even smaller than that of silicon CMOS with similar gate lengths. The small $f_{\text{max}}$ was attributed to large gate resistance, large graphene/metal contact resistances, and non-optimized designs. By contrast, the RF transistors presented here and based on C-face graphene show comparable (high) $f_\tau$ and $f_{\text{max}}$ values. A balanced $f_\tau$ and $f_{\text{max}}$ represents an optimal situation for many RF circuits. These reproducible large $f_{\text{max}}$ values result from the T-gate design, the self-alignment technique, and the inherent small contact resistance in the devices.

Due to the T-gate geometry, the gate resistance in this work is on the order of 1 $\Omega/\mu$m of gate width, an order of magnitude smaller than that without a T-gate, and over two orders of magnitude smaller than the highly doped nanowire gates \(^{24, 28, 31}\).

The superior performance of our GFETs is also in part due to our self-alignment process that is different from that shown in previous reports \(^{28, 31, 32}\). The source and drain contacts are aligned close to the gate foot, instead of being aligned to the gate head. The alignment accuracy (10 nm) is determined by the accuracy in length and the edge roughness of the gate head. This self-alignment method minimizes the ungated portion of the graphene channel, thereby reducing the access resistance and increasing the maximum current density. Parasitic effects that could arise from the contact metal under the overhanging gate head \(^{28}\) are minimized with a high gate foot (70nm). The parasitic capacitance induced by the T-gate is therefore no more than 10% the gate capacitance of the device made with high-$\kappa$ dielectric (excluding probe pads and connections), even for 50nm gate lengths.

Contact resistances $R_c$ are also limiting factor for high frequency performance. In previously reported high frequency GFETs, the $R_c$ ranges from ~300 $\Omega\mu$m to several $k\Omega\mu$m \(^2\). While the effect of $R_c$ on $f_\tau$ has been extensively discussed in the context of GFET, its effects on $f_{\text{max}}$ is often ignored, even though large $R_c$ degrades $f_{\text{max}}$ more than $f_\tau$. This is because with large $R_c$ and small channel resistance most of the input power is dissipated by the contacts instead of the channel. The absence of current saturation in graphene exacerbates this effect \(^{32}\), since in the case of GFET there is no mechanism for the channel resistance to become much larger than the contact resistance. In contrast for $f_\tau$, by driving the GFET at higher $V_{\text{ds}}$, a peak transconductance can be achieved, similar to that in small $R_c$ devices.

The small contact resistance $R_c$ between the Pd/Au layer and C-face epitaxial graphene is the key to the record high $f_{\text{max}}$ (both with and without de-embedding) reported here. Since $R_c < 100$ $\Omega\mu$m, at its best operation point, most of the RF input power can be amplified in the graphene channel itself. Following ref \(^3\), $f_{\text{max}}$ can be calculated from
\[ f_{\text{max}} = \frac{f_T}{2 \sqrt{g_D (R_G + R_{SD})} + 2 \pi f_T R_G C_G}, \]

where the channel conductance \( g_D \approx 30 \text{mS} \), source-drain access resistance \( R_{SD} \approx 15 \Omega \) (due to the self-alignment, the source-drain access resistance comprises of primarily contact resistance on the source and drain), gate resistance \( R_G \approx 3 \Omega \) and gate capacitance \( C_G \approx 5 \text{fF} \) are estimated for a dual 7\( \mu \text{m} \) wide T-gate, with the values discussed above. According to the above equation, we find \( f_{\text{max}} = 74 \text{ GHz} \) for the 100 nm gate length device, in remarkable agreement with the measured \( f_{\text{max}} \). Because of the low gate resistance, \( f_{\text{max}} \) is primarily determined by \( f_{\text{max}} = \frac{f_T}{2 \sqrt{g_D (R_G + R_{SD})}} \). A low contact resistance together with a low gate resistance is thus key to the comparable and high values of \( f_T \) and \( f_{\text{max}} \).

In summary, we have demonstrated a new fabrication method for RF transistors using high mobility C-face epitaxial graphene for the first time. With T-gate and self-alignment techniques to minimize gate and access resistance, record high \( f_{\text{max}} \) of 38 GHz and 70 GHz, before and after de-embedding, respectively, are observed on GFETs with 100 nm gate length. Reproducible results on multiple devices show comparably high \( f_T \) and \( f_{\text{max}} \), which is ideal for implementation in high frequency circuits. Our studies open a pathway to better power gain performance in GFETs, and show a great potential of C-face epitaxial graphene in high frequency applications.

**Acknowledgement**

This research was supported by the W. M. Keck Foundation, the AFSOR grant No. FA9550-10-1-0367, and the NSF MRSEC Program under Grant No. DMR-0820382. Zelei Guo would like to thank Thomas J. Beck for helpful discussions.
References


Figure captions

Figure 1
(a) Raman spectra of C-face monolayer graphene before (black) and after (red) subtraction of SiC Raman peaks, showing the G and 2D graphene peaks. Note that the D peak is very small. (b) Conductivity versus top gate voltage on C-face monolayer graphene 2 μm x 4 μm Hall bar. On the n-doped side, Hall mobility μ_Hall = 7500 cm²/Vs at n = 1.6x10¹² / cm² (V_g = 0V), and FET mobility is μ_FET = 8700 cm²/Vs. On the p-side, μ_FET ~5000 cm²/Vs. Inset: AFM image of a C-face monolayer graphene (scale bar 5μm) over SiC steps. The white lines are graphene pleats characteristic of C-face graphene.

Figure 2
Process flow for GFET fabrication. (a) Monolayer graphene on C-face of SiC. (b) The T-gate is patterned using tri-layer resist and e-beam lithography, followed by aluminum seed layer deposition and ALD of Al₂O₃. (c) Ti/Au is deposited on top as the gate metal. (d) After lift-off, the T-gate stands on graphene. Note the sides of the T-gate coated by an insulating layer. (e) Angle deposition of Pd/Au to form self-aligned contacts. (f) Ti/Au source and drain are deposited on top of the self-aligned contacts as probing pads.

Figure 3
(a) Scanning electron microscope (SEM) image of a T-gate with 100nm gate foot fabricated with a tri-layer resist, and source and drain contact metal aligned to the gate foot (scale bar 100nm). (b) SEM image of a dual-gate GFET on C-face SiC (scale bar 1μm). Source (S), drain (D) and gate (G) are indicated.

Figure 4
DC characteristics of GFETs on C-face graphene with 100 nm and 250 nm gate lengths. (a)-(b) Current density plotted as a function of gate voltage at V_d =-0.1V and -0.5V for gate length (a) 100nm and (b) 250nm. (c)-(d) Drain-source IV characteristics at gate voltage ranging from 0V to 3V at 0.5V step on (c) 100nm and (d) 250nm gate devices. Maximum current density of ~2.6mA/μm is observed.

Figure 5
High frequency characteristics of GFETs on C-face graphene with different gate lengths. (a) H₂₁ versus frequency on two devices with 100 nm gate length; device 1 (red) before de-embedding (Solid red line) and after de-embedding (red circles); device 2 (blue, ibidem). Cutoff frequency (by extrapolation of theoretical slope of 20dB/decade to |H₂₁|=1) for device 1 and 2 are before de-embedding f₁= 41 GHz (31GHz), resp, and after de-embedding f₁=110 GHz (90 GHz), resp. (b) |H₂₁| for a 250nm gate length GFET before (solid line; f₁= 32 GHz) and after (circles, f₁= 60 GHz) de-embedding. (c) Mason’s unilateral gain versus frequency for the two 100 nm gate GFETs shown in (a). Intercepts U₁²=1 (slope 20dB/dec) give f_max =38 GHz and 33 GHz before and 70 GHz after de-embedding. (d) Same as (c) for the 250nm gate length GFET in (b) f_max=36 GHz and 58 GHz before and after de-embedding.