Charge carrier velocity in graphene field-effect transistors

Marlene Bonmann,a) Andrei Vorobiev, Michael A. Andersson, and Jan Stake

Chalmers University of Technology, Terahertz and Millimetre Wave Laboratory, Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-41296 Gothenburg, Sweden

(Received 7 September 2017; accepted 19 November 2017; published online 5 December 2017)

To extend the frequency range of transistors into the terahertz domain, new transistor technologies, materials, and device concepts must be continuously developed. The quality of the interface between the involved materials is a highly critical factor. The presence of impurities can degrade device performance and reliability. In this paper, we present a method that allows the study of the charge carrier velocity in a field-effect transistor vs impurity levels. The charge carrier velocity is found using high-frequency scattering parameter measurements followed by delay time analysis. The limiting factors of the saturation velocity and the effect of impurities are then analysed by applying analytical models of the field-dependent and phonon-limited carrier velocity. As an example, this method is applied to a top-gated graphene field-effect transistor (GFET). We find that the extracted saturation velocity is ca. $1.4 \times 10^7$ cm/s and is mainly limited by silicon oxide substrate phonons. Within the considered range of residual charge carrier concentrations, charged impurities do not limit the saturation velocity directly by the phonon mechanism. Instead, the impurities act as traps that emit charge carriers at high fields, preventing the current from saturation and thus limiting power gain of the GFETs. The method described in this work helps to better understand the influence of impurities and clarifies methods of further transistor development. High quality interfaces are required to achieve current saturation via velocity saturation in GFETs. © 2017 Author(s).

All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.1063/1.5003684

The frequency range of electronic components is continuously being pushed towards higher frequencies. Particular interest is focused on the terahertz domain, due to the potential applications in imaging for medicine, spectroscopy, and wireless communication.

To increase the performance of high-frequency transistors regarding the figures of merit, namely, the transit frequency ($f_{T}$) and the maximum frequency of oscillation ($f_{\text{max}}$), new materials and technologies are still being explored. A critical factor is the interface between different materials. Impurities degrade the performance and reliability of a device. It is important to find characterization methods to study the origin of impurities and to understand how they affect device performance.

The saturation velocity of charge carriers in a transistor channel at high fields is an important material property for achieving high $f_{T}$ and $f_{\text{max}}$ values. Because of the large intrinsic carrier mobility and intrinsic saturation velocity of graphene, many efforts have been made to apply this fairly new material in high-frequency transistors. Therefore, in the following paragraph, graphene field-effect transistors (GFETs) are considered.

In previous investigations on the charge carrier velocity in GFETs using dc drain current measurements, pulsed current-voltage (I-V) measurements or transit frequencies were employed. The dc drain current method does not separate velocity and concentration, and hence, the evaluated velocity is affected by trapping/de-trapping. The pulsed I-V method avoids the slow trapping mechanisms and measures intrinsic velocity; hence, it does not allow the effects of impurities to be studied. Furthermore, the rapid (nanosecond) pulses drive charge carriers on a time scale that is much faster than that on which the energy coupling to the adjacent gate and substrate dielectrics can occur. This deviates from the velocity saturation effects in GFET structures. Under real application conditions, the saturation velocity is believed to be limited by intrinsic graphene optical phonons (OPs), surface optical phonons (SOPs) in the dielectrics, and self-heating. The transit frequency method, as published, provides velocity only, i.e., without concentration.

In this work, a method is introduced to study the charge carrier velocity in top-gated chemical vapour deposited (CVD) GFETs without the need to use a pulsed I-V measurement technique, thus allowing the study of charge carrier velocity under real application conditions. We analyze GFETs with typical top-gate design developed for high frequency applications since it allows for direct association of the evaluated carrier velocity with the GFET design and material features and their further development. Microwave measurements of high-frequency scattering parameters (S-parameters) and dc I-V characteristics are combined to determine the charge carrier velocity and charge carrier concentration independently. This allows us to demonstrate how the carrier generation from traps limits the drain current saturation. Transit frequencies are calculated from the S-parameters and are used to find the velocity of the charge carriers directly from the transit time via delay time analysis. Knowing the velocity of the charge carriers allows us to find

---

a)marbonnm@chalmers.se
the concentration of the charge carriers from the dc I-V characteristics, which is used in the analysis of phonon-limited saturation velocity. Through the use of the proposed method, the limiting factors of the saturation velocity in top-gated GFETs can be analysed. Furthermore, the effects of impurities on the mobility, the saturation velocity, and the current saturation are investigated.

Details and sequences of the fabrication steps and the characterization are given in the supplementary material. Sets of GFETs with gate length \( L = 1.0 \mu m \) and different total gate width \( W \) values of 2.5, 5, 10, and 20 \( \mu m \) are fabricated and characterized. In the analysis below, if not mentioned specifically, the GFETs with total gate width \( W = 20 \mu m \) are used. The variation in the concentration of impurities between samples is inherent to the fabrication process and is used to study the effect of impurities on the charge carrier velocity in the GFETs. Figure 1 shows a typical optical micro-photo with two gate fingers connected in parallel and a schematic 3D view of the gate stack of the GFETs.

The general outline of the developed method of analysis of the charge carrier velocity in GFETs is as follows. The transfer characteristics are used to extract the residual charge carrier concentration, the low-field mobility, and the contact resistance. Transit frequencies are calculated from the S-parameters \(^{18} \) and are then converted into delay times using resistance. Transit frequencies are calculated from the carrier concentration, the low-field mobility, and the contact transfer characteristics are used to extract the residual charge carrier concentration from the output characteristics; finally, the charge carrier concentration is used in the analysis of phonon-limited saturation velocity vs impurity levels.

Typical output and transfer characteristics of a GFET are presented in the supplementary material. Saturation of the drain current is not apparent. Fitting of a commonly used semi-empirical model \(^{19} \) to the measured drain resistance vs gate voltage \( R(V_g) \) is used to extract the mobility of the charge carriers \( \mu_0 \), the residual carrier concentration \( n_0 \), and the contact resistance \( R_c \), which includes the resistance of the metal-graphene transfer regions and the access resistance of the ungated regions. This fitting is possible under the presumption that Coulomb scattering dominates and the mobility does not depend on the concentration of the charge carriers.\(^{20} \) The value for \( n_0 \) determines the broadening of the fitting curve at the Dirac point \( V_{Dir} \). See supplementary material for a detailed discussion of the limitations of this fitting approach.

The velocity of charge carriers in field-effect transistors is evaluated using delay time analysis.\(^{21-23} \) The total current delay through the device is expressed as

\[
\tau_{tot} = \frac{1}{2\pi f_{T,ext}} = \tau_{int} + \tau_{ext} + \tau_{pad}
\]

\[
= \tau_{int} \left(1 + \frac{R_c}{R - R_c}\right) + \frac{C_g \cdot W \cdot L}{2} \frac{C_{pad}}{g_{m,ext}W}.
\]

where \( f_{T,ext} \) is the transit frequency calculated from the measured S-parameters, \( \tau_{int} \) is the transit delay, \( \tau_{ext} \) is the delay time required to charge the parasitic parts of the active device region, \( \tau_{pad} \) is the delay time associated with charging the gate pad capacitance \( C_{pad} \), and \( g_{m,ext} \) is the extrinsic trans-conductance normalized per unit width. Since the GFETs operate in the linear regime, we assume that \( C_g = C_{gs} = C_g \cdot W \cdot L/2 \), where \( C_g \) and \( C_{gs} \) are the gate-source and gate-drain capacitances, respectively, and \( C_g \) is the gate capacitance per unit area.\(^{24} \) Here, we neglect the fringing field effect. \( \tau_{pad} \) is de-embedded by extrapolating the dependence \( \tau_{tot}(1/W) \) to \( 1/W = 0. \)\(^{25} \) With the aim of de-embedding and finding \( \tau_{pad} \), we have selected and measured a set of GFETs with different widths but similar \( n_0 \) values in the subsequent analysis. Knowing \( C, R, R_c, \) and \( \tau_{pad} \), the intrinsic transit delay \( \tau_{int} \) is then calculated using Eq. (1). Finally, the intrinsic transit time is used to calculate the intrinsic transit frequency

\[
f_{T,int} = \frac{1}{2\pi \tau_{int}}.
\]

Figure 2(a) shows the extrinsic transit frequency \( f_{T,ext} \) vs the drain voltage \( V_d \) and the intrinsic transit frequency \( f_{T,int} \) vs the intrinsic electric field in the channel

\[
E_{int} = -\frac{V_d}{L} \left(1 - \frac{R_c}{R}\right)
\]

for devices with \( n_0 = (1.7, 1.9, 2.8) \times 10^{12} \) cm\(^{-2} \) (circles, squares, and diamonds). The extrinsic delay and the pad delay are responsible for the reduced \( f_{T,ext} \) compared to \( f_{T,int} \). Additionally, a larger \( n_0 \) seems to decrease the transit frequency.

Under the condition of \( V_{d,int} < V_{d,sat} = |V_g - V_{Dir}| + en/C \), where \( V_{d,int} = E_{int} \times L \) and \( n \) is the charge carrier concentration, we assume that the current regime is unipolar and that the velocity of the charge carriers and the field-dependent mobility are calculated as

\[
v = \frac{L}{\tau_{int}} \quad \text{and} \quad \mu_T = \frac{L}{\tau_{int} E_{int}}.
\]

The velocity of the charge carriers, which is calculated using Eq. (4), is presented in Fig. 2(b) for the device with \( n_0 = 1.7 \times 10^{12} \) cm\(^{-2} \). It can be seen that the velocity saturates for the fields above \( E_{int} = 1–1.5 \) V/\( \mu \)m, which corresponds well to the results reported in Ref. 12. To evaluate the saturation velocity \( (v_{sat}) \), we fitted the carrier velocity calculated from the transit delay using Eq. (4) with an analytical expression of the field-dependent carrier velocity in the range \( E_{int} = 0–1.3 \) V/\( \mu \)m to avoid the effect of self-heating\(^{12,26} \).
which assumes that estimate the saturation velocity using an analytic model, 
that a good fit is achieved, which validates the method of 
the Coulomb scattering depends on the charged impurity
concentration \(n\). The mobility increases with lower residual carrier con-
be inversely proportional to the residual carrier concentra-
\[ v \propto \frac{1}{n} \]

FIG. 2. (a) Intrinsic transit frequency vs electric field in the channel for devices with \(n_0 = (1.7, 1.9, 2.8) \times 10^{12} \text{ cm}^{-2} \) (circles, squares, and diamonds) at \(V_g = -2 \text{ V} \). The extrinsic transit frequency vs drain voltage for the device with \(n_0 = 1.7 \times 10^{12} \text{ cm}^{-2} \) is indicated in the same graph by open circles. Dashed lines are polynomial fitting curves and serve as a guide to the eye. (b) The carrier velocity for the device with \(n_0 = 1.7 \times 10^{12} \text{ cm}^{-2} \) was calculated using Eq. (4) and fitted by the empirical expression of Eq. (5) (solid line) using \(\gamma = 3\), \(\mu_0 = 1920 \text{ cm}^2/\text{V s}\), and \(v_{\text{sat}} = 1.4 \times 10^7 \text{ cm/s vs the electric field in the channel. The effective saturation velocities calculated using Eq. (8) for graphene with Al2O3 OPs (dotted), graphene with SiO2 OPs (dashed), and graphene with SiO2 and Al2O3 OPs are also shown (dashed-dotted).}

\[ v = \frac{\mu_0 E_{\text{int}}}{\left[1 + (\mu_0 E_{\text{int}}/v_{\text{sat}})^2\right]^{1/2}}, \quad (5) \]

where \(\mu_0\) is the low-field mobility and \(\gamma\) is a fitting parameter. The fitting result using \(\gamma = 3\), \(\mu_0 = 1920 \text{ cm}^2/\text{V s}\), and \(v_{\text{sat}} = 1.4 \times 10^7 \text{ cm/s is shown in Fig. 2(b). It can be seen that a good fit is achieved, which validates the method of evaluating the carrier velocity from the transit time.}

The low-field mobility found from the delay time analysis \((\mu_0)\), together with the low-field mobility found by fitting the drain resistance model \((\mu_R)\) vs the residual concentration of charge carriers found by fitting the commonly used drain resistance model, is shown in the supplementary material. The dependencies of both \(\mu_0\) and \(\mu_R\) can be approximated to be inversely proportional to the residual carrier concentration. The mobility increases with lower residual carrier concentrations because of reduced scattering. According to the self-consistent theory by Adam et al., \(n_0\) the mobility limited by the Coulomb scattering depends on the charged impurity concentration \((n_{\text{imp}})\), which directly defines the residual concentration \((n_0)\) of charge carriers. However, the absolute values of \(\mu_0\) are about two times larger. This can be explained by underestimation of \(\mu_0\) caused by the effect of oxide traps on the channel transport characteristics. We estimate the saturation velocity using an analytic model, which assumes that \(v_{\text{sat}}\) is limited by inelastic emission of OPs and can be described as

\[ v_{\text{sat}}(n, T) = \frac{2 \omega_{\text{OP}}}{\pi \sqrt{\pi n}} \left[1 - \frac{\omega_{\text{OP}}^2}{4\pi n e_F} N_{\text{OP}} + 1\right], \quad (6) \]

where \(\omega_{\text{OP}}\) is the OP energy and \(N_{\text{OP}} = 1/[\exp(\hbar \omega_{\text{OP}}/k_B T) - 1]\) is the phonon occupation. Since the channel is unipolar, the charge carrier concentration is calculated as

\[ n = \frac{L}{W e \mu_0 R - R_c}. \quad (7) \]

We assume that an effective saturation velocity \((v_{\text{sat}}^{\text{eff}})\) defined by several different OP mechanisms and Matthiessen’s rule \(n_{\text{OP}}\) in terms of velocity can be applied as

\[ \frac{1}{v_{\text{sat}}^{\text{eff}}} = \frac{1}{v_{\text{sat}}^{\text{G}}} + \frac{1}{v_{\text{sat}}^{\text{SiO2}}} + \frac{1}{v_{\text{sat}}^{\text{Al2O3}}} + \frac{1}{v_{\text{sat}}^{\text{SiO2}},} \quad (8) \]

where \(v_{\text{sat}}^{\text{G}}\) is the saturation velocity limited by the graphene zone-edge OPs \((\hbar \omega_{\text{G}} = 200 \text{ meV})\), \(v_{\text{sat}}^{\text{SiO2}}\) and \(v_{\text{sat}}^{\text{Al2O3}}\) are the saturation velocities limited by the surface OPs of the SiO2 substrate \((\hbar \omega_{\text{SiO2}} = 55 \text{ meV})\) and the Al2O3 gate dielectric \((\hbar \omega_{\text{Al2O3}} = 87 \text{ meV})\), and \(v_{\text{sat}}^{\text{si}}\) is the saturation velocity associated with OPs of the impurities. Self-heating is not taken into account since we do not see a reduction in drain current. With only the graphene OPs, the model gives \(v_{\text{sat}}^{\text{G}} = 5.1 \times 10^7 \text{ cm/s vs the measured v}_{\text{sat}}\). Figure 2(b) shows the effective saturation velocities calculated using Eqs. (6) and (8) for three different OP combinations: (i) graphene with Al2O3, (ii) graphene with SiO2, and (iii) graphene with SiO2 and Al2O3 OPs. The charge carrier concentration present in Eq. (6) is calculated using Eq. (7), leading to \(n = 2.1 \times 10^{12} \text{ cm}^{-2}\), which is found from the GFET output characteristic with \(\mu_T = 1000 \text{ cm}^2/\text{V s}\) from Eq. (4) corresponding to \(E_{\text{in}} = 1.3 \text{ V/um}\). Clearly, the combination of the graphene and Al2O3 OPs overestimates and the combination of the graphene, SiO2, and Al2O3 OPs underestimates the measured \(v_{\text{sat}}\), whereas the combination of the graphene and the SiO2 OPs only gives a good agreement. The lack of contribution of the Al2O3 OPs can be explained by the formation of the SiO2 layer at the top interface during the growth of the GFET gate dielectric at elevated temperatures, as it was shown that, due to the low diffusion barrier, the Si atoms can move almost freely on the graphene layer. We assume that the Si atoms have access to the top dielectric/graphene interface at the edges of the dielectric/graphene mesa and/or via pores in graphene, since our CVD graphene coverage is estimated to be below 90%. Figure 3(a) shows the saturation velocities calculated from the transit delay (circles) using Eqs. (4) and (5) for different GFETs vs residual carrier concentration. It can be seen that there is a clear correlation, i.e., \(v_{\text{sat}}\) decreases with \(n_0\). This could be explained by the additional contribution from OPs associated with impurities, which is represented by the last term in Eq. (8). However, when taking different charge carrier concentrations \(n\) into account for the GFETs calculated with Eq. (7) using the measured output characteristics and using these \(n\) values in Eq. (6) to estimate \(v_{\text{sat}}\), we can obtain the normalized dependence \(v_{\text{sat}}^{\text{norm}} = v_{\text{sat}}/v_{\text{sat}, \text{G}}\) [Fig. 3(a), squares]. In this case, there is no more apparent dependence on \(n_0\). Concerning phonon scattering, the effective saturation velocity is not directly dependent on \(n_0\) within the range of residual charge carrier concentrations considered in this work. Theoretically, the charge carrier concentration in
the channel is determined by the applied gate voltage and drain voltage. Comparing the charge carrier concentration for different devices at $|V_g - V_{Dir}| = 4.7$ V, far away from the Dirac point to avoid the influence of puddles, and $E_{int} = 1.5$ V/µm indicates that the decrease in $v_{sat}$ with $n_0$ can possibly be explained by the emission of electrons from traps (impurities) at high fields. Larger $n_0$ values correlate with a higher trap concentration and lead to additional charge carriers [Fig. 3(b)], which, according to Eq. (6), decreases the saturation velocity.

In conclusion, we presented a method for the analysis of the charge carrier velocity in the transistor channel and demonstrated its application using the example of top-gated CVD GFETs with different impurity concentrations. In general, the proposed method can be applied for field-effect CVD GFETs with different impurity concentrations. In this work, the GFET with the lowest residual carrier concentration ($n_0$), i.e., the lowest impurity level ($n_{imp}$) of $n_0 = 1.7 \times 10^{12}$ cm$^{-2}$ afford a saturation velocity $v_{sat} = 1.4 \times 10^7$ cm/s and an intrinsic transit frequency $f_{T,int} = 22$ GHz at a gate length of 1 µm. Analysis using a model based on optical phonon scattering at the two interfaces in the vicinity of graphene indicates that at this impurity concentration, the effective $v_{sat}^{eff}$ is limited mainly by the SiO$_2$ and graphene OPs. However, the impurity concentration and the accompanying emission of charge carriers at high fields are too high to allow saturation of the drain current to be achieved. Employing technological processes that result in a reduction in the impurity concentration and hence a reduction in the emission of charge carriers at high fields is a possible way to achieve drain current saturation via velocity saturation and thus obtain higher power gain of the GFETs. In addition, replacing the SiO$_2$ substrate and Al$_2$O$_3$ top dielectric with materials with higher OP energies, e.g., sandwiching graphene between hexagonal boron nitride, allows us to increase the saturation velocity up to $3 \times 10^7$ cm/s or $5 \times 10^7$ cm/s and the intrinsic transit frequency up to 48 GHz or 80 GHz at a gate length of 1 µm.

See supplementary material for a detailed description and discussion of device fabrication and characterization.

This work was supported in part by the Swedish Research Council (VR) under Grant No. 2014-5470, in part by the Swedish Foundation for Strategic Research (SSF) under Grant No. SE13-0061, and in part by the EU Graphene Flagship Core 1 Project.


