Radio-Frequency Linearity of Carbon-Based Nanoscale Field-Effect Transistors

by

Ahsan Ul Alam

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Solid State Electronics

Department of Electrical and Computer Engineering
University of Alberta

Abstract

This Ph.D. thesis focuses on the radio-frequency (RF) linearity of carbon-nanotube field-effect transistors (CNFETs) and graphene field-effect transistors (GFETs). The thesis can be categorized into three stages.

In the first stage, the RF linearity potential of CNFETs has been investigated by considering an array-based device structure under the first approximation of ballistic transport. A nonlinear equivalent circuit for ballistic field-effect transistors is used to compare the linearity of CNFETs to conventional MOSFETs. It is shown that nanotube devices working at high frequencies are not inherently linear, as recently suggested in the literature, and that CNFETs exhibit overall linearity that is comparable to their MOSFET counterparts. The nonlinear quantum capacitance is identified to be a major source of high-frequency nonlinearity in CNFETs. The impact of device parameters such as oxide capacitance, channel width, and tube pitch are also investigated.

In the second stage, a modified top-of-the-barrier model (MTBM) capable of simulating ballistic transport in GFETs is developed. The model captures band-to-band (Klein-Zener) tunneling, which is important in zero-bandgap materials, and it accounts for variations in the densities of states between the channel and the source and drain regions. The model is benchmarked against a sophisticated solver (based on the non-equilibrium Green's function approach) and is shown to have very good quantitative agreement. The utility of the modified TBM is demonstrated by investigating and comparing the RF linearity of GFETs to that of CNFETs and conventional MOSFETs.

In the third stage, the RF linearity potential of short-channel GFETs is assessed by using the modified top-of-the-barrier approach developed in stage 2, again under the first approximation of ballistic transport. An intrinsic GFET is examined to reveal the key features of GFET linearity, and extrinsic parasitics are then included to assess the overall RF linearity. It is shown that short-channel GFETs can be expected to have a signature behavior versus gate bias that includes a constant-linearity region at low gate bias, sweet spots of high linearity before and after the gate bias for peak unity-current-gain frequency, and poor linearity at the gate bias corresponding to the peak unity-current-gain frequency. It is otherwise found that a GFET offers overall linearity that is comparable to a conventional MOSFET and a CNFET, with the exception that the amount of intermodulation distortion in a GFET is dominated by the drain-injected carriers, a unique outcome of graphene's lack of a bandgap.

To

The three most special ladies of my life,

My mother, my wife, and my daughter

And

My life-long role model

My father

Acknowledgements

I would like to begin my acknowledgement by thanking the Almighty for giving me the opportunity, will, and provisions to pursue and complete this Ph.D.

I would like to express my heartfelt gratitude to my supervisor, Dr. Mani Vaidyanathan. He has not only been a remarkable supervisor to me but also a wonderful role model. In the span of these last six years, he has helped me grow as a self-sufficient, confident, and detail-oriented professional. His passion for quality work and his ever-positive approach to solve problems has guided me through the ups and downs of my Ph.D.

I would also like to take this opportunity to thank Dr. Diego Kienle of Universität Bayreuth, Germany, for being a wonderful mentor and friend throughout my Ph.D. I am deeply indebted to him for his guidance and especially for all those long emails, beautifully explaining the answers to my endless queries. Special thanks to Dr. Prasad Gudem of Qualcomm Inc. for his valuable feedback on my work on GFET linearity.

I am grateful to the members of my supervisory and Ph.D. examination committee, Dr. Sandipan Pramanik, Dr. Masum Hossain, Dr. Ray DeCorby, Dr. Dileepan Joseph, and Dr. Xihua Wang for their comments and suggestions that were invaluable in elevating the quality of this thesis.

I am thankful to my past and present colleagues of the Nanoelectronics Research Laboratory, Dr. Navid Paydavosi, Dr. Sabbir Ahmed, Kyle Holland, and Michael Wong. I will always cherish all our enlightening and passionate discussions at the whiteboard.

I am deeply indebted to Rezu Kaka, Salima Chachi, Sakin, and Lamisa, who have become my family away from home throughout the course of my Ph.D. I would also like to thank my dear friend Saif and Afsana for always being there for me.

I can only offer my deepest gratitude and regards to my loving parents Nusrat Sultana and Md. Badrul Alam. Their affection, dedication, and endless sacrifice have made me the person I am today.

Finally, I would like to acknowledge the immeasurable support of my loving wife Rezwana. For the last 6 years, she has singlehandedly taken care of me, my daughter Areeba, the whole household, and her own graduate research, which has enabled me to focus on my work. This Ph.D. would not have been possible without her sacrifice and tremendous support.

Special thanks to the F.S. Chia Ph.D. Scholarship, Queen Elizabeth II Graduate Scholarship, the Natural Sciences and Engineering Research Council (NSERC) of Canada, and AWR Corp. for the wonderful financial and technical support during this work.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1. Motivation

For more than half a century, Moore's law has ruled the development of electronics. However, as we approach the year 2020, the exponential growth promised by Moore's law is about to come to an end [1-4]. This realization has started a massive search for alternatives to current silicon technology. The hope is that a new material with better electrical properties will continue to provide improved performance, or that a new material will provide a path to new applications, even if the desired (exponential) scaling of silicon is no longer possible.

An array of exotic carbon-based materials, including one-dimensional (1D) carbon nanotubes and two-dimensional (2D) graphene, have shown great promise in replacing or augmenting silicon for future electronics. The advantages of these 1D and 2D forms of carbon over conventional silicon are multifold. First, the reported carrier mobility in both carbon nanotubes and graphene are much higher than in state-of-the-art silicon [5, 6], suggesting the possibility of generally improved device behavior even at long channel lengths (where mobility is a valid figure of merit for transport). Second, consistent with the first point, the electron mean-free paths in both nanotubes and graphene are much larger than in silicon (100 nm in nanotubes and graphene [7] vs. 40 nm in silicon [8]), which means that scattering will not have as deleterious an effect on device performance

at short channel lengths. Third, the 1D and 2D nature of these materials facilitate excellent electrostatics for superior gate control at smaller technology nodes.

Consistent with the above observations, both carbon-nanotube field-effect transistors (CNFETs) and graphene field-effect transistors (GFETs) have been reported to offer the potential for excellent high-frequency figures of merit that could meet or exceed the specifications of the International Technology Roadmap for Semiconductors (ITRS) [9], including unity-current-gain frequencies (f_T) and unity-power-gain frequencies ($f_{\max}$) in the range of THz for channel lengths in the 20-nm range [10, 11]. Given the potential for such high cutoff frequencies, combined with their unique electrical properties, carbon nanotubes and graphene have garnered particular attention as promising candidates for *analog* high-frequency, or radio-frequency[1] (RF), electronics.

One of the key figures-of-merit for RF applications is linearity[2], which measures the degree of distortion generated by the nonlinear mixing of an input signal with jammers. High linearity is desirable in wireless communication systems, where signals can get distorted due to interference and intermodulation from nearby communication bands and channels (the so-called "jammers"). Advantages in linearity would hence provide a specific motivation to further pursue new materials for RF applications, over and above simply trying to extend Moore's law.

Linearity is difficult to ascertain *directly* from first-principles physical descriptions, such as a non-equilibrium Green's functions (NEGF) approach or the Boltzmann transport equation (BTE). Rather, a compact circuit model is needed. In this thesis, we use such a model, based on the "top-of-the-barrier" approach[3] [12-14]. We calibrate the

[1] The terms "radio-frequency" and "high-frequency" are used interchangeably in this thesis.

[2] Ideally, we desire a small ac input signal to be *linearly* amplified by a transistor, without interference from jamming signals that are nearby in frequency. However, the nonlinear nature of a transistor's current- and charge-voltage behavior will cause the output signal to be a *distorted* version of the input, due to the nonlinear mixing of a desired input with nearby jammers. *Linearity* is the property --- defined formally by a figure of merit known as the "third-order input-intercept point (IIP3)" --- that characterizes the extent of such distortion.

[3] The "top-of-the-barrier" model is a method of describing charge and current in any ballistic (collisionless) transistor. Historically, it is based on an original formulation by Natori [12] that was refined by Rahman *et al.* [13] and which was then used to create an equivalent circuit (for purely linear operation) according to Hasan *et al.* [14]. The key elements of the model are provided in Appendix A. Unfamiliar readers may wish to review Appendix A prior to reading Chapters 2 and 3. For the discussion in this introductory

model with NEGF or BTE results, and employ it to study the RF linearity offered by emerging channel materials, particularly those based on carbon.

1.2. This Work

Many studies have been conducted on the dc (static) and *linear* high-frequency behavior of carbon-based transistors, while their *nonlinear* high-frequency behavior remains relatively unexplored. This work is thus concerned with the *nonlinear* high-frequency performance of emerging carbon-based, field-effect transistors (FETs). The aim of the research was to accomplish the following tasks:

(1) Use a nonlinear circuit model for carbon-nanotube field-effect transistors, based on the top-of-the-barrier approach [12-14] and calibrated with BTE simulations, to shed insight into their RF linearity.

(2) Develop a nonlinear circuit model for graphene field-effect transistors based on an extended top-of-the-barrier approach to include effects arising from the positional variation of the density of states and band-to-band tunneling, both of which arise from graphene's unique lack of a bandgap.

(3) Use the developed nonlinear circuit model of GFETs from (2) to shed insight into their RF linearity.

To accomplish the above tasks, the Ph.D. research can be categorized into three stages. All three stages are complete, and have led to a significant journal paper in the *IEEE Transactions on Nanotechnology* [15] and a conference paper at the *IEEE SISPAD* conference held in Glasgow, Scotland in September, 2013 [16]. A third manuscript has been submitted to the *IEEE Transactions on Microwave Theory and Techniques* in August, 2014 [17].

---

chapter, it suffices to note that under the assumption that the electrons traveling from the source to the drain experience *no* collisions, the charge and current in the transistor can be known by keeping track of what happens at a *single point* in the channel, *i.e.*, the point at which the conduction band reaches its highest energy, the so-called *top of the* source-to-drain energy *barrier.*

1.3. Stages of Work

A summary and a description of the key points of each stage of the Ph.D. are provided below; interested readers can find more details in Chapters 2, 3, and 4.

1. **RF Linearity Potential of Carbon-Nanotube Transistors Versus MOSFETs**

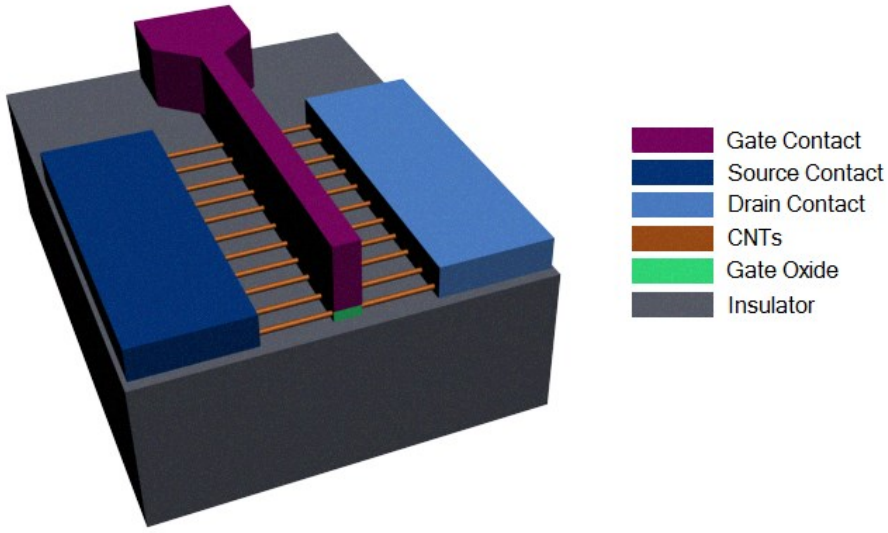
Summary:

CNFETs are one of the most promising alternatives to conventional silicon MOSFETs, and their RF figures of merit show great potential for applications in high-frequency electronics [18-21]. To date, there has been a limited amount of theoretical and experimental work done on the *nonlinear* RF performance of CNFETs [22-24]. In the first stage of the Ph.D., we hence use a compact, nonlinear, small-signal equivalent circuit for array-based CNFETs based on the top-of-the-barrier approach [12-14] to investigate CNFET linearity. With this model, the linearity of CNFETs is shown to be comparable to --- *not better than*, as previously reported [22] --- their conventional MOSFET counterparts. Key mechanisms defining the nonlinear behavior of both CNFETs and MOSFETs are identified. Our work also reveals that a smaller pitch in CNFET arrays yields not only greater drive current but also better linearity (provided that the devices are sufficiently wide). Our published work [15] on this stage of the Ph.D. was reviewed as being *"an excellent paper,"* with the study described as *"systematic"* and *"very well-organized,"* and the conclusions deemed to be *"of great interests* [sic] *to researchers working on carbon nanotube electronics."*

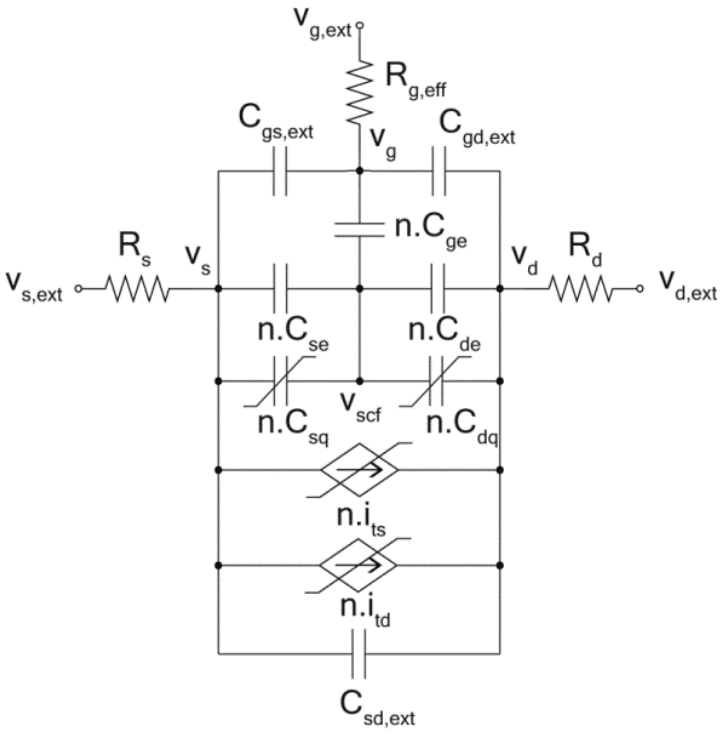
Key Points:

Linearity is an extremely important figure of merit in RF electronics because it dictates the ability of an analog amplifier to faithfully amplify a small ac input signal. In 2007, Baumgardner *et al.* [22] claimed that CNFETs have the potential to achieve "inherent linearity," *i.e.*, *perfectly* distortionless amplification. Although the study only considered distortion arising from the nonlinear current-voltage relationship (transport nonlinearity) of a CNFET, and neglected distortion arising from the nonlinear capacitance, the promise of "inherent linearity" was intriguing.

Hence, to probe the claim further, we investigated the impact of the nonlinear capacitance.



(a)



(b)

Fig. 1.1. (a) Schematic of the array-based CNFET used in stage 2 of the Ph.D. (b) Extrinsic nonlinear, small-signal equivalent circuit of an array-based CNFET.

The array-based CNFET used in the study is reproduced here in Fig. 1.1(a), and Fig. 1.1(b) shows the nonlinear, small-signal equivalent circuit based on the top-of-the-barrier approach, assuming the CNFET contains n tubes. Details on the circuit are available in Chapter 2, where it is presented in Fig. 2.1.

Our work reveals that although the linearity of a CNFET such as that in Fig. 1.1(a) seems to be superior to its MOSFET counterpart when only transport nonlinearity is considered, the nonlinear quantum capacitance makes the *overall* linearity of a CNFET comparable to that of a conventional MOSFET; this outcome is evident from the results displayed in Fig. 1.2. The nonlinear quantum capacitance is identified as the factor that degrades the linearity of CNFETs. The nonlinear 1D density of states of a CNFET yields a highly nonlinear quantum capacitance [the element $n \cdot C_{\mathrm{sq}}$ in Fig. 1.1(b)], which is to be contrasted with the essentially linear quantum capacitance arising from the essentially linear 2D density of states of a MOSFET.

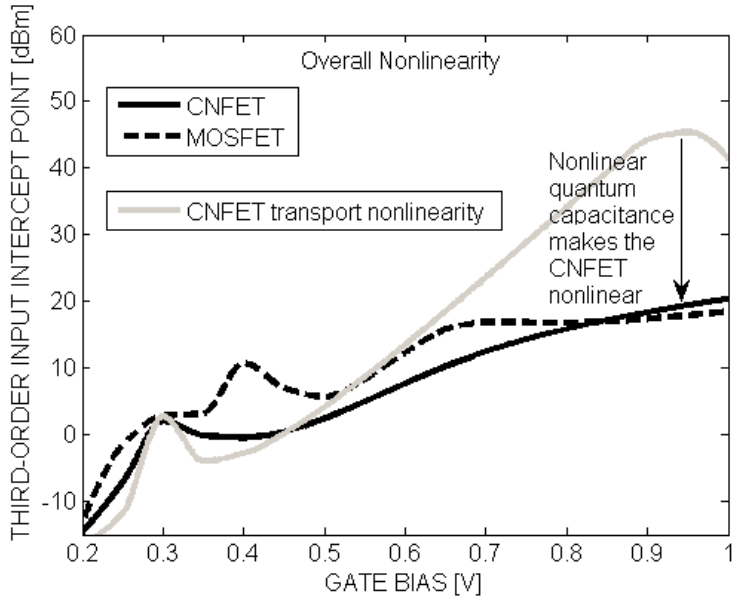


Fig. 1.2. Intrinsic (neglecting parasitics) IIP3 vs. gate bias for a 20-nm ballistic array-based CNFET and a 20-nm ballistic conventional MOSFET.

It was also claimed in [22] that CNFETs can be made "inherently linear" in the limit of infinite gate-oxide capacitance. However, our work reveals the opposite to occur, *i.e.*, that the device becomes *less linear* under the influence of a large gate-oxide capacitance. The basic result can be discerned from the results in Fig. 1.3; the details of this figure (which is the same as Fig. 2.13) are discussed in Chapter 2, and here the essence can be noted by focusing on the solid, dotted, and stippled curves. The solid curve shows the linearity with a typical gate-oxide capacitance, achieved with a gate dielectric constant ε_r set to 16 (corresponding to hafnium oxide); increasing ε_r to 160 seems to improve the linearity, but *only* if we restrict our attention to the impact of the first conduction subband in the nanotube, as shown by the dotted curve; if we additionally account for the impact of the second conduction subband, the linearity degrades, as shown by the stippled curve. A more detailed explanation, including a connection of the behavior in Fig. 1.3 to the quantum-capacitance elements in Fig. 1.1(b), is provided in Section 2.3.3(b).

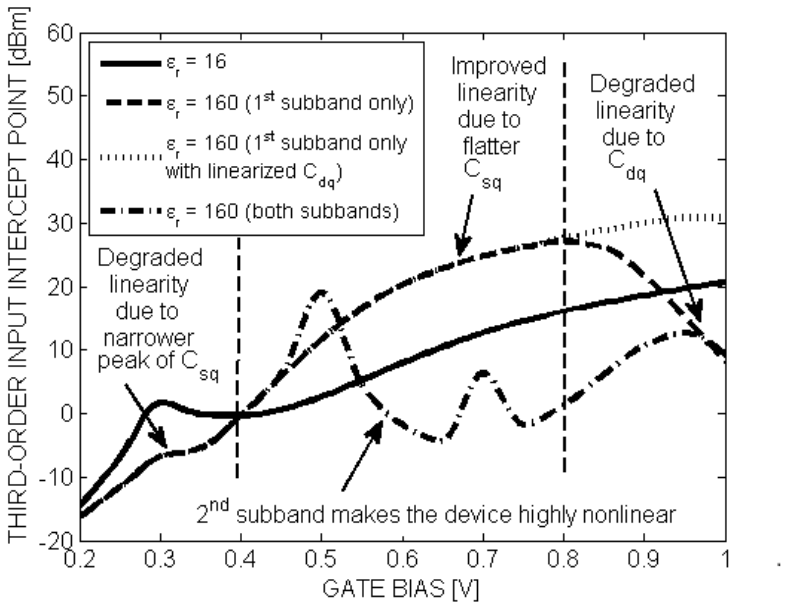


Fig. 1.3. Extrinsic (including parasitics) IIP3 vs. gate bias for a 20-nm ballistic array-based CNFET with different gate oxides.

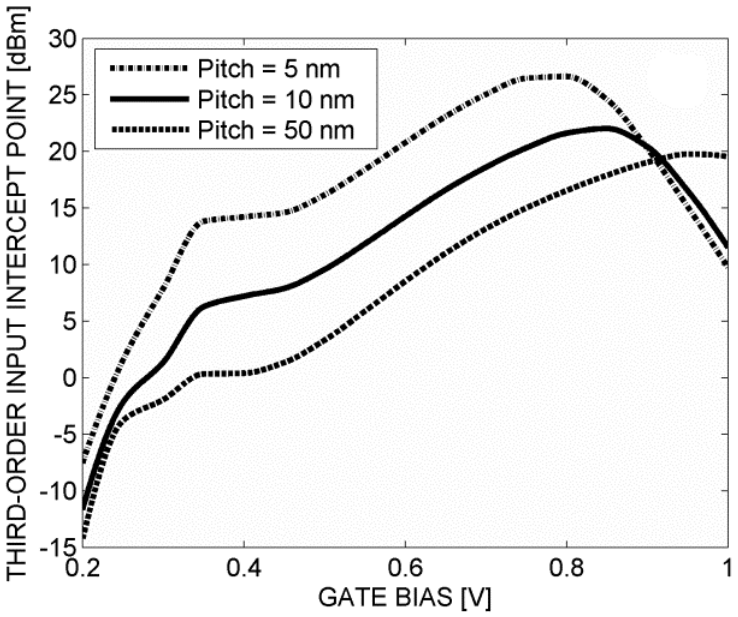


Fig. 1.4. Extrinsic (including parasitics) IIP3 vs. gate bias for a 30-μm wide ballistic array-based CNFET for different tube pitches.

Finally, in our work, we show that the extrinsic (including parasitics) linearity of an array-based CNFET depends on the tube pitch, provided that the device is sufficiently wide (≥ 30 μm). The key results are shown in Fig. 1.4. Here, for a device having a channel width of 30 μm, it is shown that the device becomes more linear as the tube pitch is decreased. This is a very favorable observation, as a smaller tube pitch also means a higher drive current and consequently a larger gain.

2. A Modified Top-of-the-Barrier Solver for Graphene and its Application to Predict RF Linearity

Summary:

Due to its unique electrical properties, graphene has recently been the center of attraction for new devices. Although the zero bandgap of graphene makes it difficult to employ in digital applications, the RF properties of graphene show great promise for high-frequency electronics [25, 26]. While GFETs have been analyzed and assessed using top-of-the-barrier models [27, 28], the accuracy of these approaches

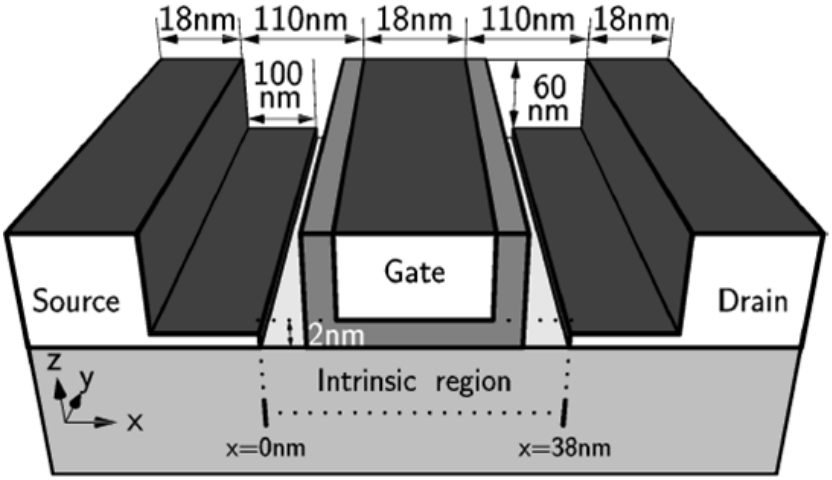
suffer when nonidealities such as the positional variation of the density of states and band-to-band tunneling become prominent. In this stage of work, we have thus modified the conventional top-of-the-barrier model to account for these nonidealities, and we have developed a compact, nonlinear, circuit model that agrees with a highly sophisticated numerical solver based on NEGF [10]. Nonlinear analysis of the circuit showed that the linearity of GFETs is very sensitive to the drain bias, unlike with conventional transistors. Our initial findings were presented at the *IEEE Simulation of Semiconductor Processes and Devices (SISPAD) 2013* [16]. Detailed analysis of the developed nonlinear model was later conducted (in the third stage of the Ph.D. discussed below) to gain further insight into the physics behind the nonlinear operation of GFETs.

Key points:

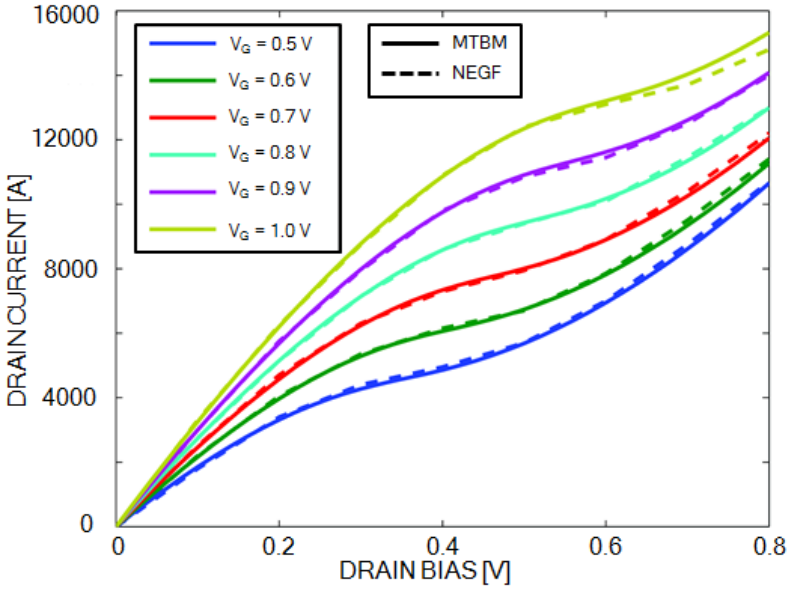
The modified top-of-the-barrier model developed in this stage captures the mentioned nonidealities through the following steps. First, we account for the fact that the densities of states in the heavily doped source and drain regions differ from the density of states (DOS) in the channel region; specifically, we construct an "effective DOS" that accounts for the fact that the DOS in the source and drain contacts vanish at the local Dirac points, precluding transmission at these energies. Second, since graphene has no bandgap, a major portion of the source-to-drain transport is mediated via band-to-band tunneling, which we include into the amended top-of-the-barrier model within the WKB approximation [29].

Fig 1.5(a) shows the structure of the GFET investigated. The structure is identical to the one used in [10]. Fig. 1.5(b) shows the current-voltage characteristics of the GFET from our amended top-of-the-barrier model and a self-consistent NEGF solver [10]. Considering the simplicity of our amended top-of-the-barrier approach, the agreement is quite remarkable. The developed model not only quantitatively matches the $i_D - v_G$ output characteristics, but also is in good agreement with more sensitive quantities involving derivatives, such as the transconductance g_m and output conductance g_d, as demonstrated in Fig. 1.6(a). The second derivatives of the current with respect to gate voltage v_G are plotted in Fig. 1.6(b), which shows an

even better improvement over the conventional model. Such improved agreement is important for an accurate prediction of the RF properties, such as linearity, which requires a proper modeling of the slope and curvature of the transistor characteristics versus voltage.

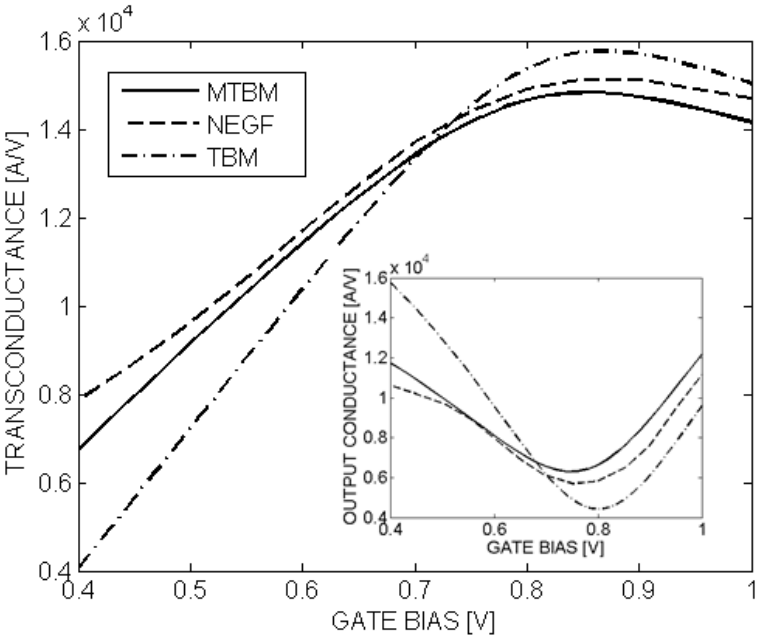


(a)

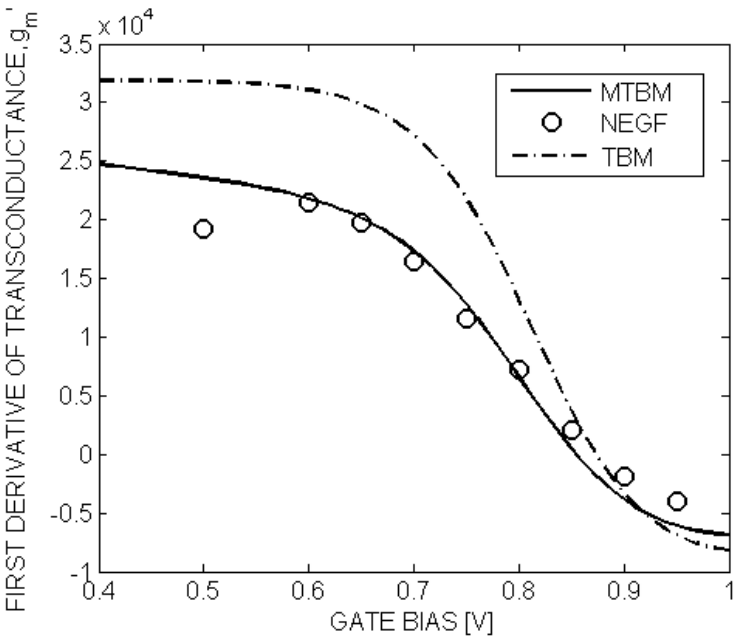


(b)

Fig. 1.5. Schematic of the simulated GFET [10]. (b) Current-voltage characteristics of the GFET from NEGF and the modified top-of-the-barrier model.

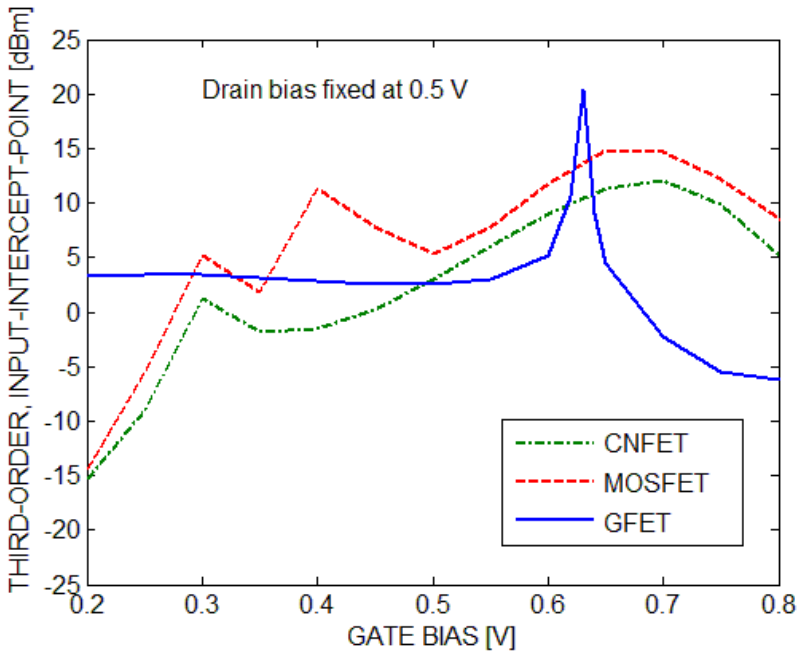


(a)

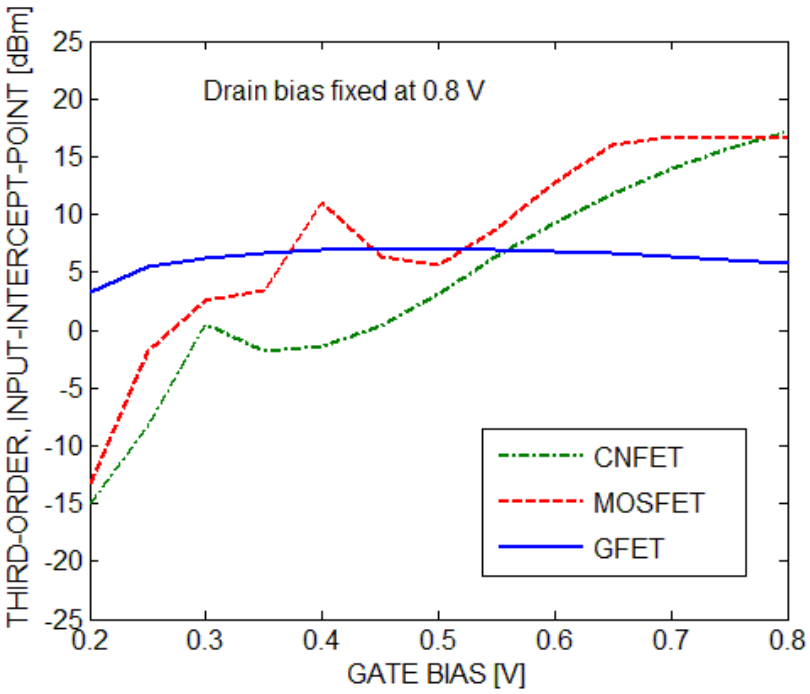


(b)

Fig. 1.6. (a) Transconductance, output conductance (inset), and (b) first derivative of transconductance with respect to gate voltage of the GFET from NEGF, conventional top-of-the-barrier model, and modified top-of-the-barrier model.



(a)



(b)

Fig. 1.7. IIP3 versus gate bias, *with the drain bias held fixed at (a) 0.5 V and (b) 0.8 V*, of the GFET compared to its CNFET and MOSFET counterparts.

The modified top-of-the-barrier model allows us to predict the circuit properties of a GFET including all key physical effects, something that would be very difficult to do directly from first-principles formalisms such as NEGF. For example, we used the modified top-of-the-barrier approach to determine the element values for a GFET's equivalent circuit, which has the form in Fig. 1.8 [similar to that in Fig. 1.1(b)]. We then used the harmonic balance (HB) solver in Microwave Office [30] to determine the GFET's IIP3.[4] Fig. 1.7(a) shows that for a drain bias of 0.5 V, the IIP3 of the GFET is comparable to its CNFET and MOSFET counterparts, where the latter have dimensions and oxide capacitances identical to the GFET. Increasing the drain bias to 0.8 V slightly improves the GFET linearity but the overall IIP3 remains comparable to that of the CNFET and the MOSFET, as shown in Fig. 1.7(b). However, unlike a CNFET and a MOSFET, where the drain biasing has a minimum effect on the characteristics of the IIP3 curve, our results show that the shape of the IIP3 curve of GFETs is quite sensitive to the drain bias.

3. RF Linearity Performance Potential of Short-Channel Graphene Field-Effect Transistors

Summary:

This stage of work exploits the amended top-of-the-barrier model developed in the second stage of the Ph.D. (discussed above) to explore GFET linearity in greater detail.

To date, there have been only a few modeling studies that explore GFET linearity [31-33], and none of them accounted for all the sources of nonlinearity relevant for RF performance, and in particular, transport *and* capacitive nonlinearities, both of which can be expected to play a role [15]. In this final stage of work, we thus provide a more detailed and comprehensive study of the RF linearity mechanisms of an 18-nm GFET, chosen for demonstration purposes and consistent with current CMOS technology nodes [9]. Based on an examination of IIP3 values for

[4] Please see footnote 2 on p. 2 for a definition of IIP3 that is sufficient for the present discussion.

intermodulation distortion under a two-tone input, our study reveals that a GFET's linearity has a *signature behavior* versus gate bias that includes a constant-linearity region at low gate bias, sweet spots of high linearity before and after the gate bias for peak f_T, and poor linearity at the gate bias corresponding to the peak f_T. We otherwise find that a GFET offers overall linearity that is comparable to a MOSFET and a CNFET, with the exception that the amount of intermodulation distortion in a GFET is dominated by the drain-injected carriers, a unique outcome of graphene's lack of a bandgap. We also examine the effects of drain bias, load resistance, and external parasitics.

Key Points:

We start our analysis by using the nonlinear small-signal circuit of Fig. 1.8 [16]. The intrinsic components of the circuit are extracted based on the modified top-of-the-barrier model (MTBM) developed in the second stage [16]. The external parasitics are then calculated with the aid of COMSOL [34] and added to obtain the complete extrinsic nonlinear circuit, an approach that was already shown [16] in the second stage to capture the nonlinear voltage dependencies of key device parameters determined from a more detailed simulator [10]. The HB solver in Microwave Office (MWO) [30] is then used to simulate the nonlinear circuit.

Fig 1.5(a) show the structure of the GFET under investigation, identical to the one used in the second stage. Fig. 1.5(b) shows the current-voltage characteristics of the GFET from our modified top-of-the-barrier model and a self-consistent NEGF solver [10]. As in the earlier stages, we assume ballistic transport, a reasonable first approximation for graphene at small channel lengths ($\lesssim 20$ nm) for the purposes of assessing performance potential, especially since the reported electron mean-free path in graphene is much larger ($\gtrsim 100$ nm) [7]. We also consider a doped MOSFET-like device, as done in recent studies to assess the performance potential of carbon-based electronics [14, 27]; short-channel MOSFET-like devices can be expected to outperform the long-channel Schottky-barrier devices prevalent today [35] and are a suitable choice to gauge performance potential.

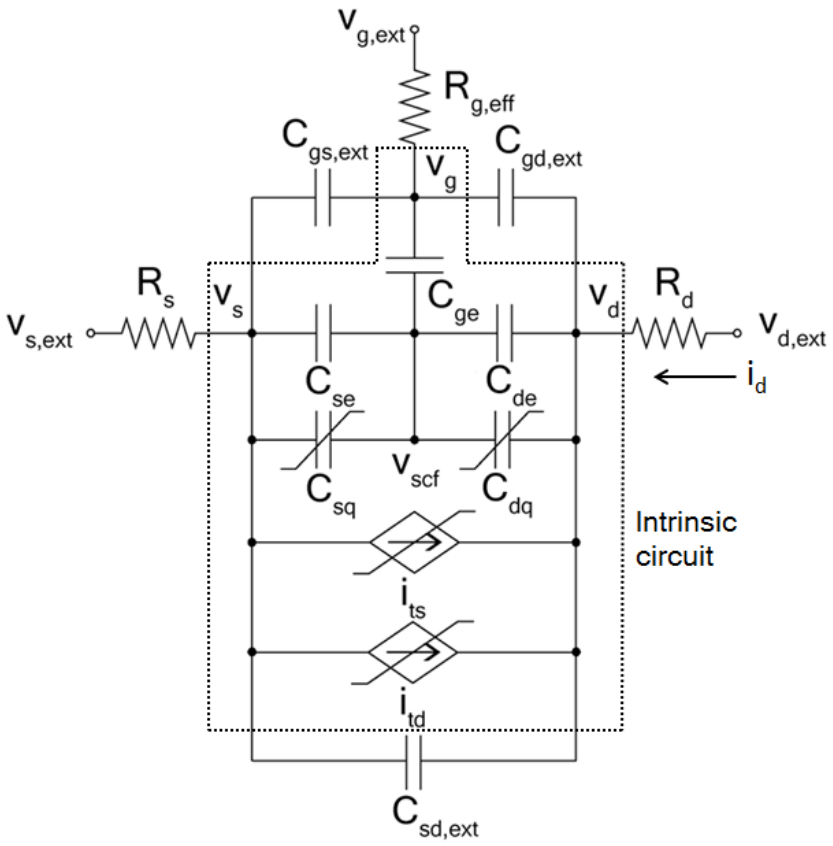


Fig. 1.8: Complete nonlinear small-signal equivalent circuit of a ballistic GFET.

To reveal the key features of GFET linearity, we first investigated the intrinsic RF linearity of a GFET, *i.e.*, the linearity determined by the dotted portion of Fig. 1.8 and excluding external parasitics. The resulting IIP3 was plotted against variations in gate bias and is shown in Fig. 1.9. The IIP3 curve has a very distinct shape (signature), with a constant linearity region (region 1), two sharp peaks at points 2 and 4, and a large dip around point 3. The presence of the peaks at points 2 and 4 means that bias sweet spots may exist where a GFET will behave very linearly. Fig. 1.9 also shows the unity-current-gain frequency f_T versus gate bias. Note that the peak f_T coincides with point 3, which means the GFET is most nonlinear at peak f_T. A detailed discussion on the mechanism behind this signature behavior of GFET IIP3 is available in Chapter 4.

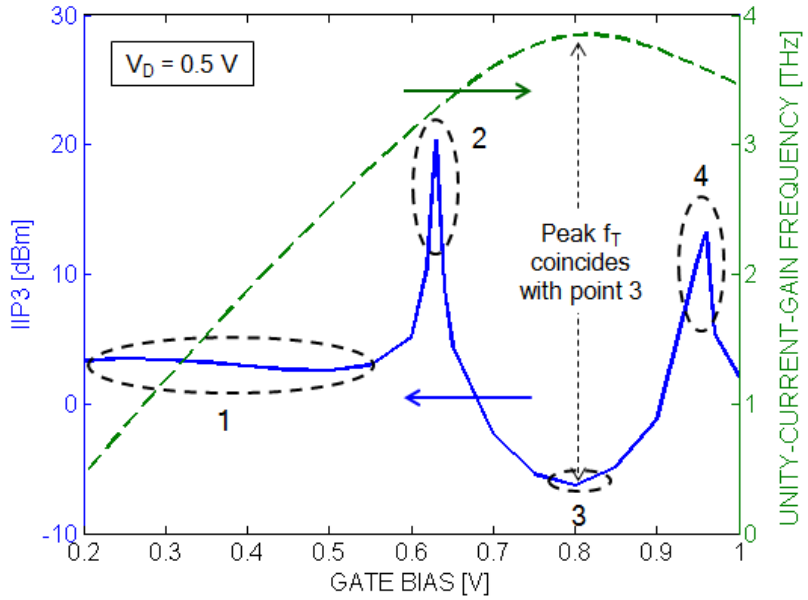


Fig. 1.9: Intrinsic IIP3 and unity-current-gain frequency f_T versus gate bias for the GFET under investigation.

A closer look at the GFET linearity mechanism reveals that the zero bandgap of graphene results in a strong drain dependence in GFET linearity. By selectively turning on and off the distortion from the components in the circuit of Fig. 1.8, it is possible to isolate the distortion originating from the source and drain. The results are shown in Fig. 1.10, which plots the GFET IIP3 due to distortion coming only from the source and drain, respectively, and compares it with the overall IIP3. For all gate biases, the linearity of the device is found to be dominated by distortion coming from the drain. This result is significantly different from a conventional field-effect transistor in which the channel material has a finite bandgap (MOSFET or CNFET), where the distortion primarily comes from the source components [15].

One obvious outcome of the drain dominance on GFET linearity is a drain-bias dependence of the overall linearity. For example, Fig. 1.11 (which is the same as Fig. 4.10) shows that in region 1 (constant linearity), a larger drain bias makes the device more linear; other aspects of the figure are discussed in Chapter 4.

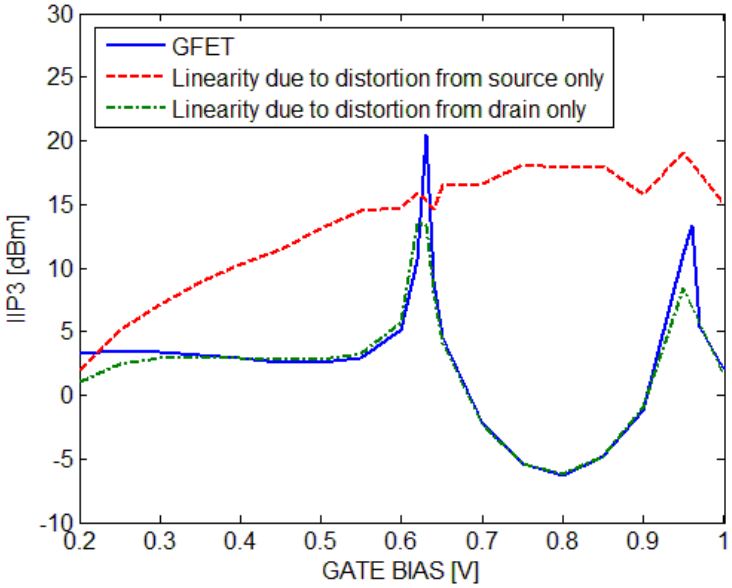


Fig. 1.10: Effect of distortion from the source and drain on IIP3.

Another outcome of the drain dominance on GFET linearity is the effect of the load resistance R_L. A larger R_L enhances the distortion from the drain components and makes the device more nonlinear whereas a smaller R_L makes the device more linear by reducing the distortion generated by the drain components. Fig. 1.12 shows the effect of R_L on IIP3, while the source resistance is held at 50 Ω. As expected, reducing the load from 50 Ω to 12.5 Ω dramatically increases the GFET IIP3 by almost 10 dB. Similarly, increasing the value of R_L degrades linearity. The unique zero bandgap of graphene (the reason behind the drain dominance) thus makes it possible to improve the linearity by reducing the load resistance. However, before reducing R_L to improve the linearity, one must consider its implications on the voltage and power gains of the device, two desirable properties of any FET operating at RF frequencies. Such considerations, along with a more detailed discussion on the drain dependence of GFET linearity and its outcomes, are provided in Chapter 4.

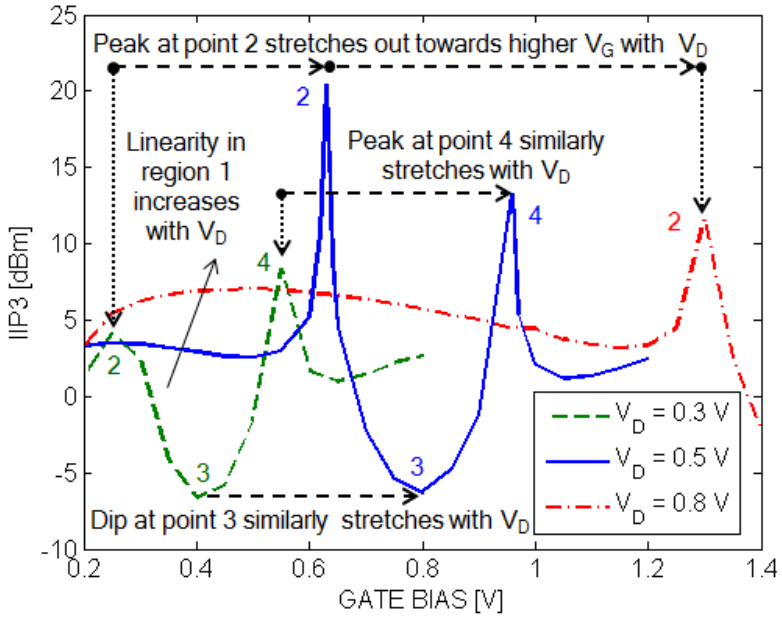


Fig. 1.11: IIP3 versus gate bias, at a few different values of drain bias.

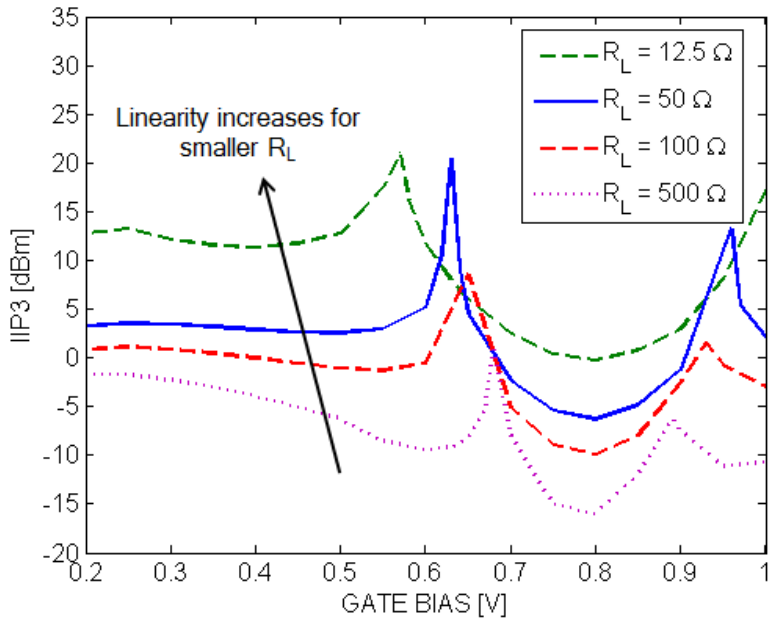


Fig. 1.12: Effect of load resistance R_L on GFET IIP3.

In order to determine if a GFET holds any promise in RF electronics in terms of linearity, we compare its linearity performance against that of a silicon MOSFET and an array-based CNFET with identical channel length ($L = 18$ nm), channel width

($W = 1$ μm), and gate capacitance. Fig. 1.13 shows that the GFET offers linearity that is, overall, comparable to its MOSFET and CNFET counterparts under this scenario. However, two differences can be flagged. First, as already discussed, the drain dependence of the GFET offers us with an opportunity to enhance its linearity by increasing the drain bias v_D or by lowering R_L, which is not possible in the other devices. Second, the GFET's linearity offers a sweet spot prior to and after peak f_T; these are the points 2 and 4 discussed earlier in conjunction with Fig. 1.9. In fact, the GFET offers its worst IIP3 at peak f_T, unlike the MOSFET and the CNFET, both of which offer their best IIP3 at peak f_T.

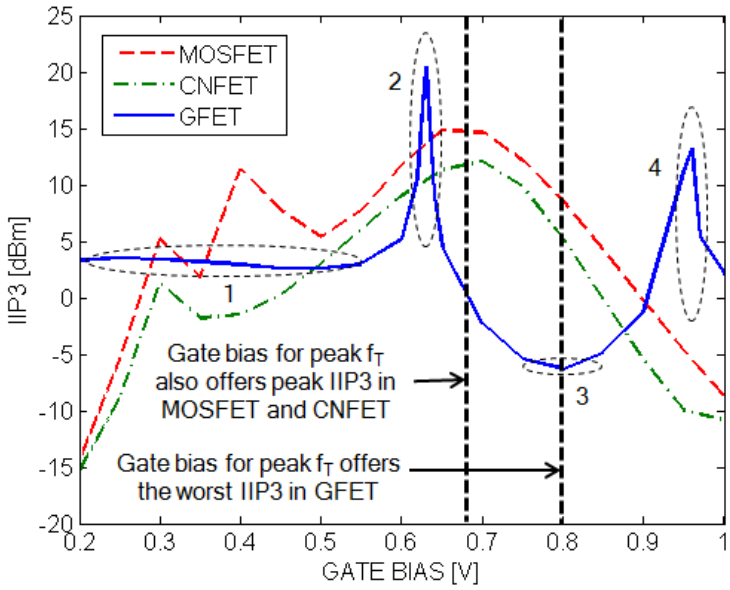


Fig. 1.13: Intrinsic linearity performance comparison of a GFET with its MOSFET and CNFET counterparts. The peak IIP3 for a CNFET and MOSFET occur at the same gate bias as peak f_T whereas for a GFET the minimum IIP3 occurs at the gate bias for peak f_T.

Simulation of the entire circuit in Fig. 1.8 reveals that the external parasitics slightly degrade the device linearity, but the signature shape identified from the intrinsic device remains, as demonstrated in Fig. 1.14. The drain contact resistance is shown to be primarily responsible for degrading the overall RF linearity of the GFET. Additionally, the presence of the drain contact resistance makes it impossible to reduce the distortion from the drain components by using a smaller R_L.

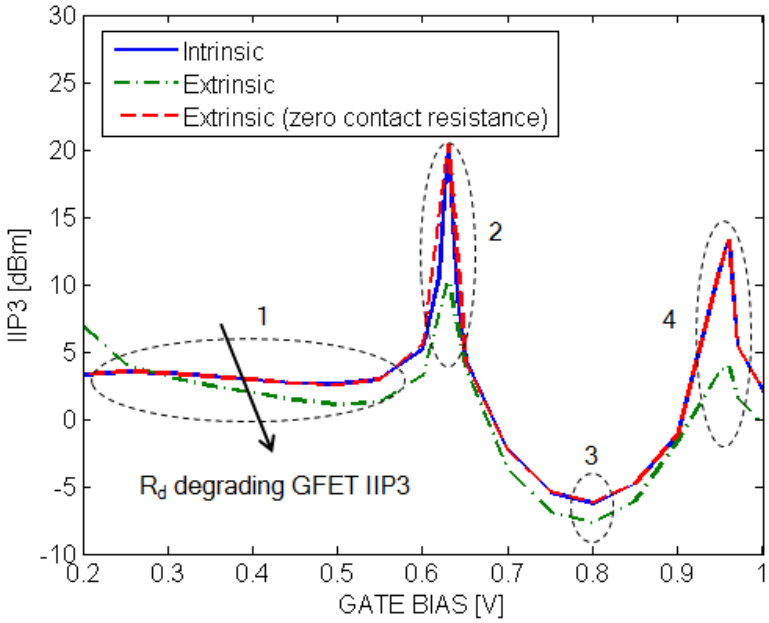


Fig. 1.14: Intrinsic and extrinsic IIP3 versus gate bias.

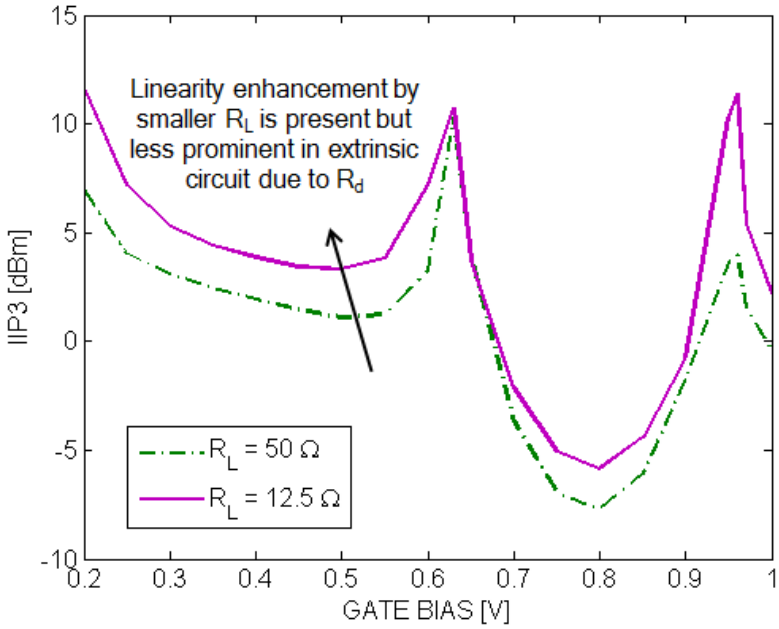
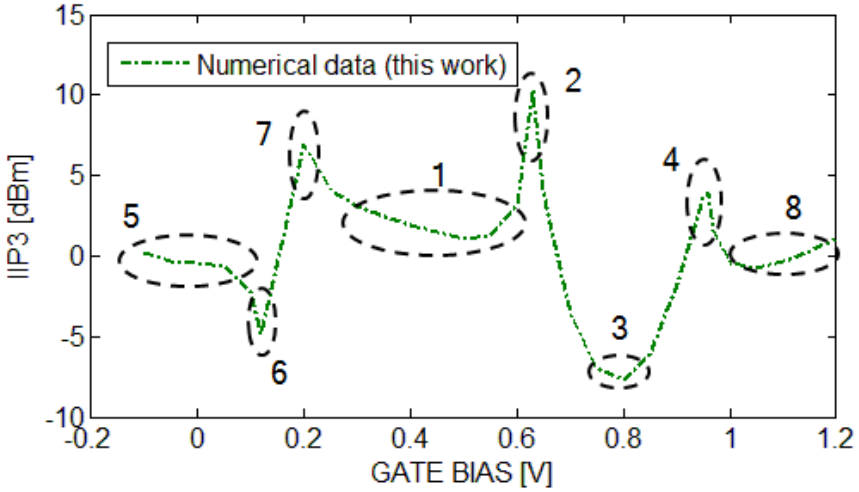


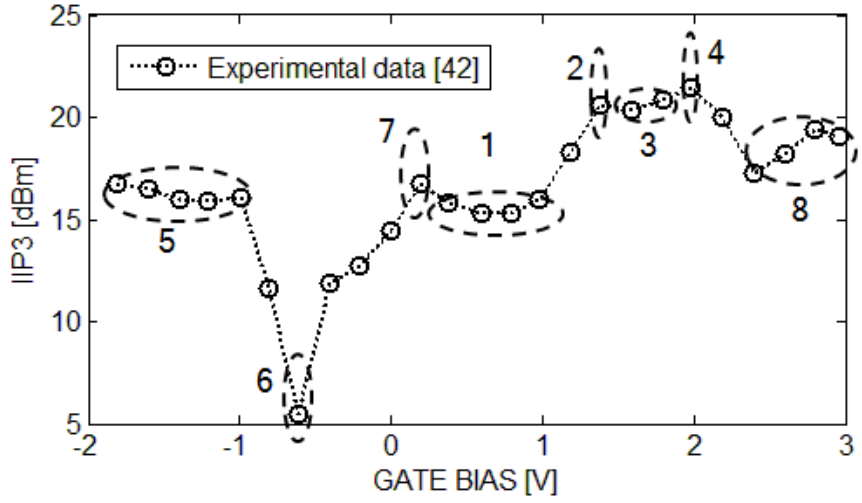
Fig. 1.15: Effect of load resistance R_L on GFET extrinsic IIP3. The improvement in IIP3 with a reduction in R_L is less pronounced than in the intrinsic case shown in Fig. 1.12.

Fig. 1.15 shows the effect of variation in R_L on linearity for the extrinsic GFET. The reduction in R_L that improved the linearity by almost 10 dB in the intrinsic

circuit, as illustrated in Fig. 1.12, only improves the linearity by 2.2 dB in the extrinsic circuit. Keeping the drain contact resistance low is hence also important to allow for potential linearity improvement by adjusting R_L.



(a)



(b)

Fig. 1.16: Qualitative comparison of (a) numerical (extrinsic) IIP3 values of the GFET under investigation with (b) experimental data [42].

Finally, we compare our IIP3 values with experimental results. To date, there has been a significant amount of experimental work done on the RF linearity of graphene transistors [36-42]. However, all of these studies were done on long-channel devices ($L \geq 250$ nm), which makes a *direct* comparison with our simulation results

impossible. Nonetheless, we can make a *qualitative* comparison to the reported IIP3 values in [42]. Fig. 1.16 shows that the key signature of the GFET IIP3 (regions 1 to 4), as identified in this work, is present even in the long-channel ($L = 700$ nm) device considered in [42]. Furthermore, by extending the gate bias values beyond the 0.2 V to 1 V range used for most of the simulations in our study,[5] we see that our model also captures the features in regions 5, 6, 7, and 8 seen in the fabricated device. For the purpose of this thesis, this comparison provides a striking qualitative validation of our modeling approach and our resulting observations on the linearity of GFETs. A more detailed discussion on the mismatch between the numerical and experimental results, and the mechanism behind the new features (regions 5 to 8), is included in Chapter 4. An even more detailed comparison and discussion will be presented in future work that is to be published separately from the work done directly for this thesis; the work done for this thesis, as summarized in this chapter and as laid out in full in Chapter 4, including the qualitative comparison to experiment, has been submitted to the *IEEE Transactions on Microwave Theory and Techniques* [17].

1.4. Summary of Contributions

The unique electrical properties of carbon-based materials have attracted a lot of attention from the semiconductor industry, and the promises of ultra-fast operation and near-ballistic transport have made carbon-based nanoscale transistors serious contenders in future RF electronics. This thesis has examined their linearity, a key RF property.

By revealing the key physical mechanisms, we have shown that the RF linearity of CNFETs and GFETs is comparable to their conventional silicon MOSFET counterpart. However, each of these devices offers unique opportunities for linearity enhancement that could potentially make them favorites in future RF applications.

[5] In comparing the numerical and experimental data in parts (a) and (b) of Fig. 1.16, the *actual* gate bias and IIP3 values are not important; these will not overlap, as the two devices involved have different channel lengths. Of relevance is the *relative* positions of the identified regions and points with respect to gate bias, and the resulting *signature* in the IIP3 behavior.

In the case of CNFETs, the low 1D density of states of the nanotubes results in a very linear transport in the device. The same 1D density of states also results in a very nonlinear capacitance, which can ultimately limit the overall linearity to be on par with conventional silicon MOSFETs. However, our work also reveals that a practical array-based CNFET can potentially offer superior RF linearity by increasing the number of tubes in the channel and thereby enhancing the linearizing feedback [p.101, 43] provided by the extrinsic circuit components.

In the case of GFETs, the RF linearity is found to be dictated by graphene's 2D density of states and its zero bandgap. The zero bandgap introduces additional distortion from drain-injected carriers, which ultimately restricts the overall RF linearity of these devices and makes them comparable to conventional silicon MOSFETs. However, the unique 2D density of states combined with the zero bandgap creates linearity sweet spots, where highly linear behavior can be expected. In addition, the zero bandgap makes the GFET linearity sensitive to the load, which offers the potential for linearity enhancement by load manipulation.

Overall, our work in this Ph.D. shows that although the RF linearity of CNFETs and GFETs are comparable to conventional MOSFETs, both offer unique opportunities for linearity enhancements. This potential for enhanced linearity combined with their superior electrical characteristics continues to make these carbon-based nanoscale transistors very promising candidates for future RF electronics.

Chapter 2

RF Linearity Potential of Carbon-Nanotube Transistors Versus MOSFETs[6]

2.1. Introduction

CNFETs are promising candidates for emerging RF electronics [18], and they have recently been predicted to offer "inherent linearity" [22]. Linearity is an extremely desirable property for transistors operating at high frequencies, particularly for wireless communications [18]; for example, a small advantage in linearity could simplify the design and implementation of low-noise amplifiers (LNAs) used in wireless receivers and hence reduce the overall production cost. Given the potentially high values of unity-current-gain frequency (f_T) [19, 21] and unity-power-gain frequency ($f_{\max}$) [20, 21] attainable by CNFETs, their potential for linear behavior merits further investigation.

To date, the study of CNFET linearity has been limited. In 2007, Baumgardner *et al.* [22] analytically showed that the current-voltage relationship of a ballistic CNFET can become "inherently linear" under certain conditions; while their approach considered only the transport current as a source of nonlinearity within the device, the claim of "inherent linearity" was nevertheless intriguing. In 2008, Curutchet *et al.* [23] measured the nonlinear behavior of a CNFET at a fundamental frequency of 600 MHz and suggested a simple model to predict the third-order intermodulation distortion. More

[6] A version of this chapter has been published [15].

recently, in 2011, Wang *et al.* [24] reported promising linearity figures of merit for a nanotube transistor working at 1 GHz. As far as we know, no further studies of the linearity of CNFETs have emerged, and a more comprehensive investigation is thus warranted.

When it comes to classifying CNFETs based on their operation, there are mainly two types: the "MOSFET-like" [44-52] and the "zero-Schottky-barrier" [51-56] structure. From an experimental viewpoint, the fabrication of the latter is easier; however, MOSFET-like structures have been predicted to outperform their zero-Schottky-barrier counterparts in both digital [57, 58] and RF [59, 60] applications. In addition, for practical implementations, it is now well-known that array-based CNFETs (vs. single tubes) must be used [18, 61]. This work thus concentrates on the linearity of array-based, MOSFET-like CNFETs.

The starting point of our analysis is to develop a simplified model for the nonlinear behavior of a ballistic single-tube CNFET based on a "top-of-the-barrier" approach [12-14]; the assumption of ballistic transport can be justified by the ongoing scaling of device size and by the aim of performing a best-case assessment. Despite being simple in nature, the model is shown to be capable of capturing the nonlinear voltage dependencies of key device properties determined from a more detailed simulator [62], such as the voltage dependencies of the drain current and the f_T. Once the intrinsic nonlinear components of the single-tube transistor are extracted, they are used to generate the intrinsic components of an array-based CNFET. The external parasitics of the array-based structure are then calculated with the aid of COMSOL [34] and are appropriately added to get the complete extrinsic nonlinear model.

The HB simulator "Microwave Office (MWO)" [30] from AWR Corp. is used to simulate the developed nonlinear model (equivalent circuit). The simulator is used to measure the third-order input-intercept point (IIP3) [63] for a two-tone input, a widely accepted linearity figure of merit. With this approach, the linearity of an array-based CNFET is investigated for both typical and limiting structures, such as a structure in which the gate electrostatic capacitance is much larger than the quantum capacitance, a

scenario that might substantially enhance the device linearity, as predicted by Baumgardner *et al.* [22].

The main contribution of this work is a comparative study of the linearity of CNFETs vs. their MOSFET counterparts. Our work reveals that CNFETs are unlikely to offer a major advantage over conventional MOSFETs in terms of RF linearity, owing to the distortion arising from the nonlinear quantum capacitance in CNFETs. We also examine the channel-width and tube-pitch dependency of CNFET linearity.

Section 2.2 presents the device structure and the development of the nonlinear equivalent circuit based on the top-of-the-barrier approach. Section 2.3 compares the linearity of CNFETs with conventional CMOS transistors of similar specifications and also probes the overall high-frequency linearity of array-based CNFETs in terms of device parameters. The conclusions of this work are summarized in Section 2.4.

2.2. Approach

2.2.1. Device Structure

Fig. 2.1(a) shows the structure of the top-gated, array-based CNFET used in this work, and Fig. 2.1(b) shows a portion of the transistor consisting of a single tube that can be considered as a single-tube CNFET. The tubes in the structure are all zigzag (16,0) carbon nanotubes with the following characteristics: a diameter of 1.25 nm; doped source and drain regions (outside the gate) that each have a background *n*-type doping level of 10^9 m^{-1} and a length of 50 nm; and an undoped *i* region (underneath the gate) of length 20 nm. The gate (or channel) length of 20 nm was chosen purely for demonstration purposes; the results will apply to any gate length for which the transport can be considered ballistic or simply as a best-case assessment. The planar gate oxide has a thickness of 2 nm and a relative dielectric constant of 16 (*i.e.*, hafnium oxide), and the gate metal has a thickness of 50 nm and a work function of 4.5 eV (*e.g.*, chrome or tungsten). The tubes sit on a thick (100-nm) layer of silicon oxide, and the total channel width is taken to be 1 μm for demonstration purposes.

Fig. 2.1(c) shows the conventional n-channel, Si MOSFET used for comparison. The gate/channel length is 20 nm and the oxide thickness is 3 nm with a relative dielectric constant of 16 (*i.e.*, hafnium oxide). The heavily doped source and drain regions have a doping density of 10^{20} cm^{-3}. The body is considered to be large enough to neglect the effect of the substrate for the sake of simplicity.

The CNFET gate capacitance is matched to that of its MOSFET counterpart by taking the number of tubes per µm to be 100, yielding a tube pitch of 10 nm. Existing literature shows that this is the maximum achievable density for CNFET arrays without significant tube-to-tube screening [64]. It is worth mentioning that a structure similar to that in Fig. 2.1(a) (but with a longer channel length and wider tube pitch) has been experimentally demonstrated [49].

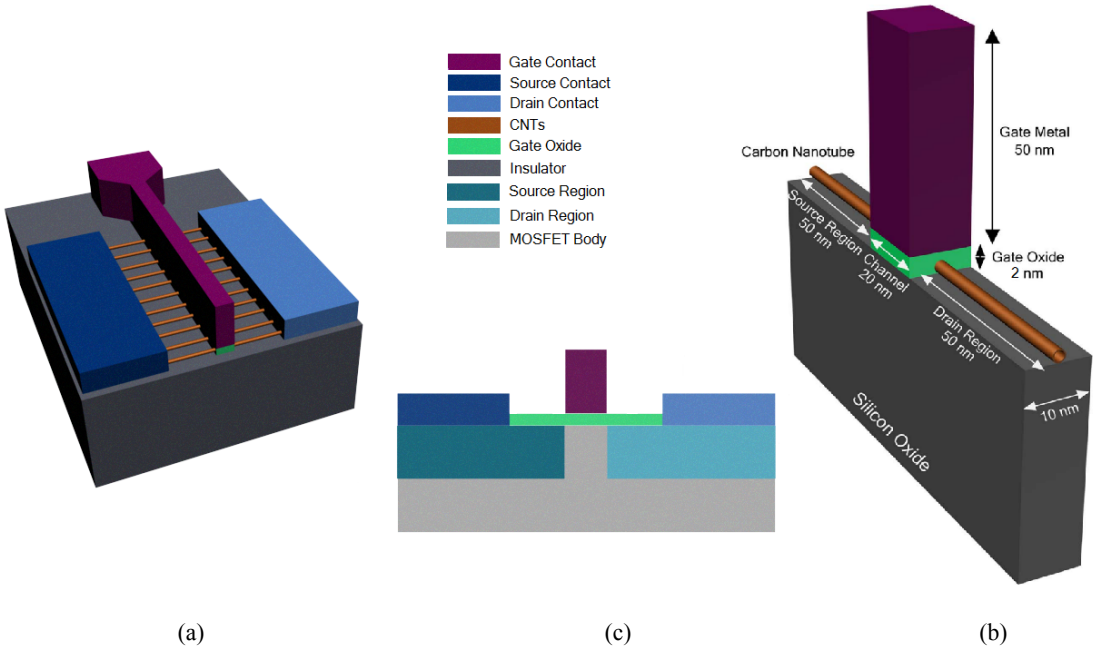


Fig. 2.1. Schematics of (a) the array-based CNFET structure, (b) a portion of the transistor consisting of a single tube that can be considered as a single-tube CNFET, and (c) the conventional MOSFET counterpart. The figures are not drawn to scale.

2.2.2. Intrinsic Equivalent Circuit

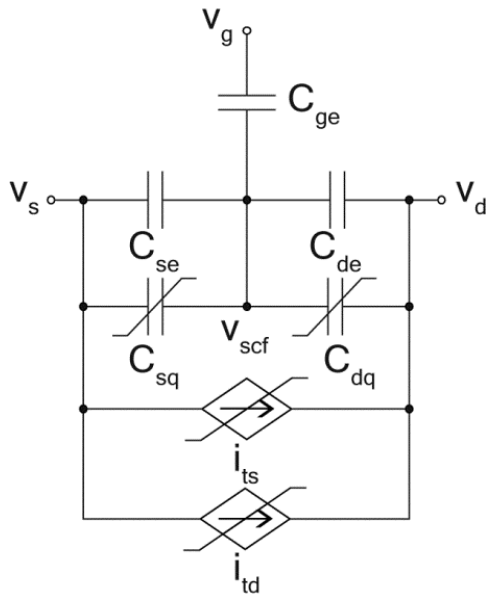


Fig. 2.2. Nonlinear equivalent circuit for a single-tube ballistic CNFET based on the purely linear topology developed in [14].

In this work, we focus on *small-signal* nonlinear operation, *i.e.*, of interest is the nonlinear behavior of small signals superimposed on dc bias values; we hence use Taylor-series expansions for all components in the small-signal equivalent circuit, where the coefficients of the series are specified by appropriate derivatives [evaluated at the bias (operating) point] of the corresponding large-signal device equations. Fig. 2.2 shows the topology of such a circuit for a single-tube CNFET, adapted from a purely linear small-signal equivalent circuit [14] based on the top-of-the-barrier model of Rahman *et al*. [13].

The elements in Fig. 2.2 are as follows: C_{ge}, C_{se}, and C_{de} are the electrostatic capacitances of the CNFET and are presumed linear, an assumption justified further below; C_{sq} and C_{dq} are the nonlinear source and drain quantum capacitances, respectively; and i_{ts} and i_{td} are nonlinear sources that model the quasi-static transport

currents[7] of the device. Based on the small-signal assumption and the relationships discussed in [13, 14], the charge-voltage or current-voltage behavior of each nonlinear element is represented by a Taylor-series expansion up to third order:

$$q_{\mathrm{sq}} = C_{\mathrm{sq1}}(v_s - v_{\mathrm{scf}}) + C_{\mathrm{sq2}}(v_s - v_{\mathrm{scf}})^2 + C_{\mathrm{sq3}}(v_s - v_{\mathrm{scf}})^3 \tag{2.1}$$

$$q_{\mathrm{dq}} = C_{\mathrm{dq1}}(v_d - v_{\mathrm{scf}}) + C_{\mathrm{dq2}}(v_d - v_{\mathrm{scf}})^2 + C_{\mathrm{dq3}}(v_d - v_{\mathrm{scf}})^3 \tag{2.2}$$

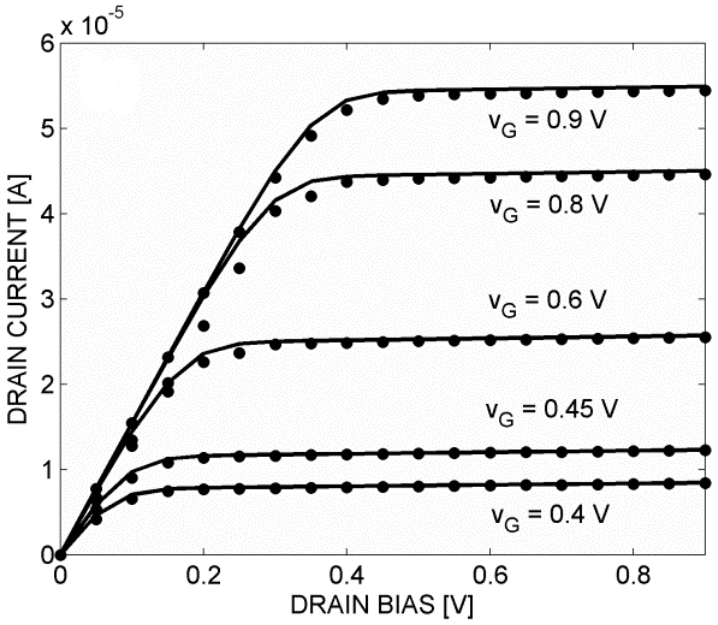
$$i_{\mathrm{ts}} = g_{\mathrm{sq1}}(v_s - v_{\mathrm{scf}}) + g_{\mathrm{sq2}}(v_s - v_{\mathrm{scf}})^2 + g_{\mathrm{sq3}}(v_s - v_{\mathrm{scf}})^3 \tag{2.3}$$

$$i_{\mathrm{td}} = g_{\mathrm{dq1}}(v_d - v_{\mathrm{scf}}) + g_{\mathrm{dq2}}(v_d - v_{\mathrm{scf}})^2 + g_{\mathrm{dq3}}(v_d - v_{\mathrm{scf}})^3 \tag{2.4}$$

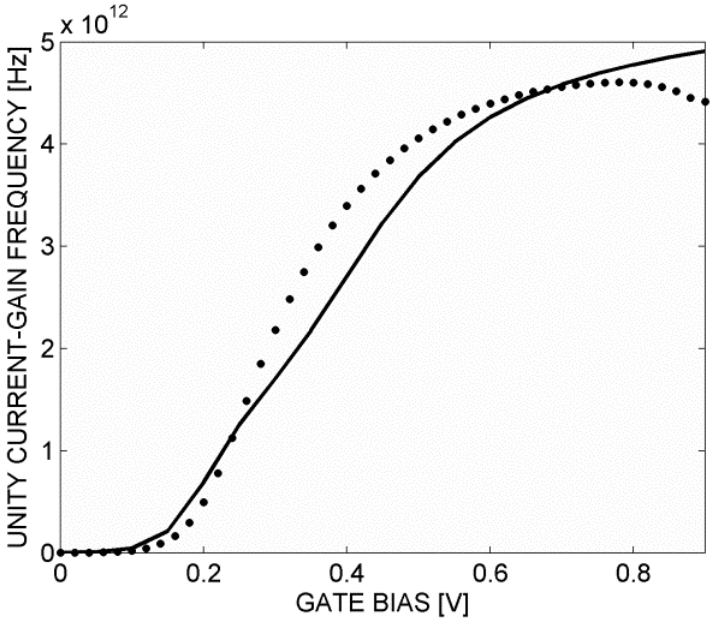
where q_{sq} and q_{dq} are the small-signal charges held by the quantum capacitances C_{sq} and C_{dq}, respectively, v_s and v_d are the small-signal source and drain voltages, respectively, and v_{scf} is the small-signal channel potential.

To extract the values of the coefficients in equations (2.1) – (2.4), the *large-signal* top-of-the-barrier equations of [13, 14] were first solved and fitted to the output of the more detailed BTE-Poisson solver reported in [62]; with the recipe provided in [13], the values of the relevant fitting parameters from this process were $E_F = -0.25$ eV, $C_{\mathrm{ge}}/C_e = 0.87$, and $C_{\mathrm{de}}/C_e = 0.01$, where $C_e \equiv C_{\mathrm{ge}} + C_{\mathrm{se}} + C_{\mathrm{de}}$ is the total electrostatic capacitance. Figs. 2.3(a) and 2.3(b) show the resulting agreement between the large-signal equations and the numerical solver [62]. With this agreement established, the coefficients in (2.1) – (2.4) could easily be obtained from the large-signal equations at any operating point of interest. The agreement in Fig. 2.3(a) shows that the nonlinear current-voltage behavior will be properly modeled by this procedure. Similarly, the agreement in Fig. 2.3(b), while more crude, shows that the overall nonlinear charge-voltage behavior of the device will be properly modeled [65] to an extent sufficient for the purposes of this work, *i.e.*, for a preliminary linearity assessment; it also shows that the agreement can be obtained under the assumption of linear electrostatic capacitances.

[7] In this thesis, we use the following convention: small-signal voltages and currents are denoted by lowercase letters with lowercase subscripts whereas total (large-signal) voltages and currents are denoted by lowercase letters with uppercase subscripts. For convenience, bias (dc) voltages and currents, which are equal to the large-signal values at the operating point, are referenced by the same notation as the large-signal quantities, with the meaning clear from the context.



(a)



(b)

Fig. 2.3. (a) Family of current vs. bias voltage curves and (b) unity-current-gain frequency f_T vs. gate bias for a 20-nm single-tube CNFET. The solid lines represent the values from the top-of-the-barrier approach [12-14] and the black dots represent data from the BTE-Poisson solver reported in [62].

2.2.3. Extrinsic Equivalent Circuit

a) Array-Based Structure

To model the behavior of an array-based structure (as opposed to a single tube), we assume all the individual single-tube CNFETs of an array have identical bias points; this can be justified by the negligible dc voltage drop on the metal gate (since the dc gate current is zero) and by the assumption that the source and drain biases are applied to all the tubes from the sides [from the left and right in Fig. 2.1(a)]. Since the tubes of an array are hence essentially in parallel, the intrinsic equivalent circuit for an n-tube array can be obtained simply by multiplying the power-series coefficients for each of the elements in Fig. 2.2 by n.

b) Parasitics

In practical CNFETs, the extrinsic (parasitic) circuit elements due to the metallic contacts are the main reason for the degradation of the cutoff frequencies (f_T and $f_{\max}$) from the theoretically predicted values. The impact of these parasitics on the linearity should also be assessed. Fig. 2.4 shows a cross-section of an array-based structure with the extrinsic capacitances marked, along with the dimensions of the gate and source/drain regions. The dashed portion of the figure (excluding the labeled extrinsic capacitances) can be modeled by the circuit of Fig. 2.2 with the element values multiplied by n, as already discussed. To such a circuit, we then add the labeled extrinsic capacitances, along with the contact resistances $R_{\mathrm{g,eff}}$, R_s, and R_d of the gate, drain, and source, respectively; all these parasitics were found as described in [21] with the aid of COMSOL [34] and by using the contact dimensions specified further below in Section 2.3.3(a). The final circuit is shown in Fig. 2.5, where v_s , v_d, and v_g are the internal node voltages on the CNTs of an array and $v_{\mathrm{s,ext}}$, $v_{\mathrm{d,ext}}$, and $v_{\mathrm{g,ext}}$ represent the external terminal voltages of the overall device. Table 2.1 in Section 2.3.3(a) lists all the component values (both intrinsic and extrinsic) for the device.

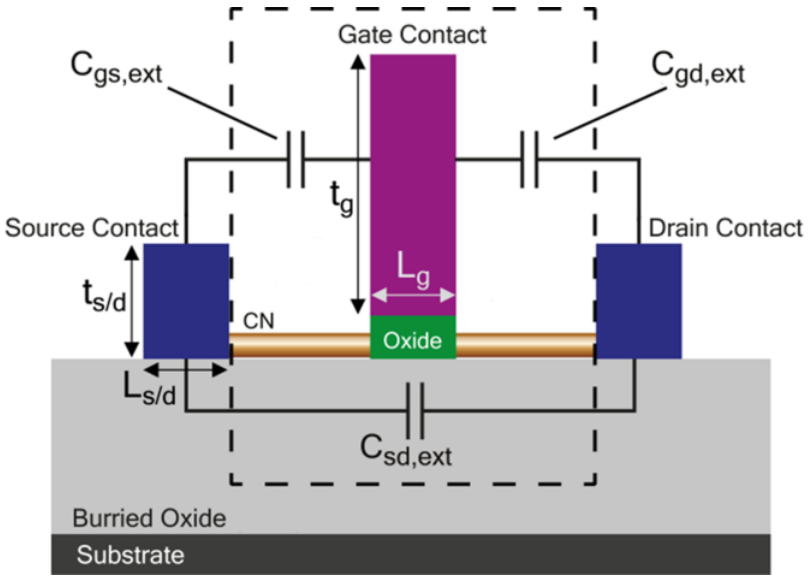


Fig. 2.4. Cross-section of the array-based CNFET structure used in this work.

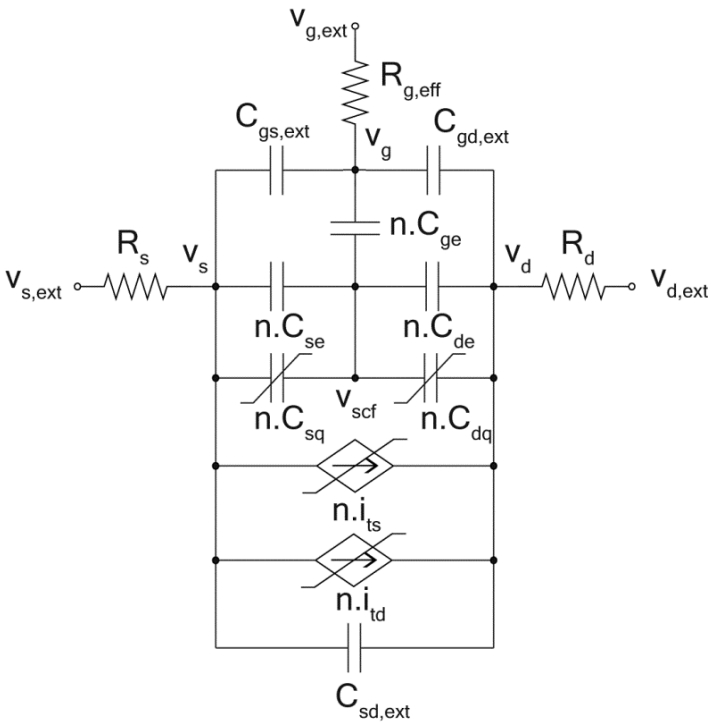


Fig. 2.5. Complete nonlinear small-signal equivalent circuit for an array-based CNFET.

2.3. Results and discussion

2.3.1. Condition for "Inherent Linearity"

In order to investigate the potential for "inherent linearity" in CNFETs, we need to establish the condition under which the *large-signal* transport current $i_{\mathrm{T}} \equiv i_{\mathrm{TS}} + i_{\mathrm{TD}}$ becomes a linear function of the gate voltage v_G, where i_{TS} and i_{TD} refer to the source- and drain-injected components of the current, respectively, and where the source is taken as the reference ($v_S \equiv 0$ V). As shown in Appendix B, under the assumption of a high drain voltage, where $i_{\mathrm{T}} \approx i_{\mathrm{TS}}$, the condition of Baumgardner *et al.* [22] for inherent linearity is equivalent to requiring that the ratio of channel charge to gate-oxide capacitance be much smaller than the gate voltage:

$$\left|\frac{\lambda}{C_{\mathrm{ge}}}\right| \ll v_G \tag{2.5}$$

where λ represents the channel charge. Based on condition (2.5), a MOSFET-like CNFET might hold promise for highly linear behavior, since the one-dimensional (1-D) density of states of a nanotube will tend to keep the channel charge λ small for a given v_G, while the use of high-k dielectrics can result in a large gate-oxide capacitance C_{ge}. However, (2.5) is derived while considering only the nonlinearity due to the transport current $i_T \approx i_{\mathrm{TS}}$, *i.e.*, it neglects distortion arising from the quantum capacitances C_{sq} and C_{dq}; this is equivalent to retaining the nonlinear small-signal current source i_{ts} in Fig. 2.2 (with $i_{\mathrm{td}} \approx 0$), while neglecting all distortion due to C_{sq} and C_{dq}. Hence, the *overall* nonlinearity of CNFETs, even if (2.5) is well satisfied, remains unclear, and it must be tested and compared to other devices.

2.3.2. CNFET vs. CMOS Transistor

a) Basis and Method for Comparison

In order to determine whether a CNFET offers any advantage over its conventional MOSFET counterpart in terms of RF linearity, we also investigate the linearity of the MOSFET shown in Fig. 2.1(c) using the same top-of-the-barrier [12-14] approach used

for the CNFET. Besides having identical gate capacitances, the MOSFET is also assumed to have identical dimensions to the CNFET, including the channel length and width, and the size and orientation of the gate, source, and drain contacts. The barrier height at the top-of-the-barrier under equilibrium is kept identical to the CNFET at 0.25 eV.

MWO was used to simulate the nonlinear circuits of the CNFET and MOSFET using the harmonic-balance (HB) technique; the circuit of Fig. 2.2 was used for both devices, with the component values replaced by total values in the case of the MOSFET and single-tube values multiplied by the number of tubes n in the case of the CNFET. The load impedance was set to 50 Ω, the usual characteristic impedance for RF applications. A two-tone source with an impedance of 50 Ω and an operating frequency of 24 GHz --- which is an application frequency of interest as identified by the 2012 ITRS [9] --- and a difference of 100 kHz between the tones were used. The source was grounded, the drain bias v_D was fixed at 0.8 V, and the gate bias v_G was varied from 0.2 to 1.0 V.

b) Nonlinearity Due Only to Transport Current

To investigate the effect of the nonlinear transport current, both transistors were simulated with the nonlinear current sources i_{ts} and i_{td} in Fig. 2.2 fully active but with linearized quantum capacitances C_{sq} and C_{dq}, *i.e.,* with all the coefficients in (2.3) and (2.4) retained but with the higher order coefficients in (2.1) and (2.2) set to zero. The third-order input-intercept point (IIP3) at different gate biases is shown in Fig. 2.6.

The basic shape of the IIP3 vs. gate bias is comparable for the two devices, including the presence of an IIP3 sweet-spot in the low-bias region and an improvement in IIP3 with increasing bias; more importantly, as shown, with only the transport-current nonlinearity, the CNFET has a significantly higher IIP3 at high bias.

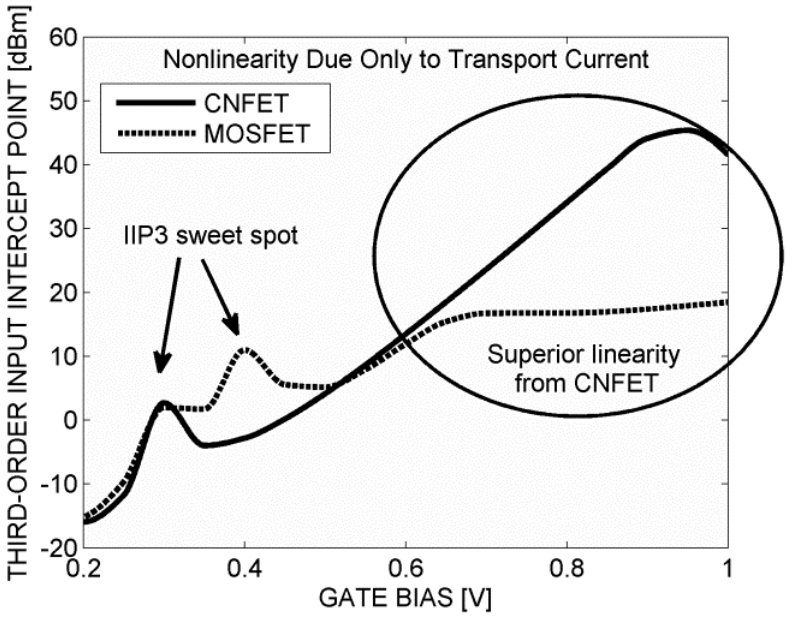


Fig. 2.6. Intrinsic IIP3 vs. gate bias for a 20-nm ballistic array-based CNFET and ballistic MOSFET, considering only nonlinear transport current.

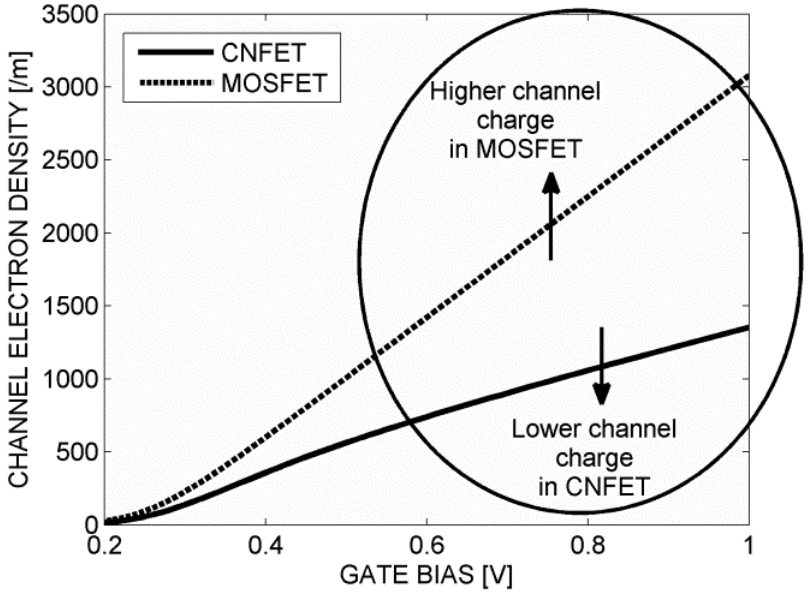


Fig. 2.7. Channel charge (electron) density of the 20-nm ballistic array-based CNFET and ballistic MOSFET.

The reason for the better linearity of the CNFET at high bias in Fig. 2.6 can be discerned from the curves in Fig. 2.7, where it is shown that the CNFET has a

significantly lower channel charge at high gate bias ($v_G \geq 0.5$ V), causing (2.5) to be better satisfied. To examine this outcome in greater detail, we first compare the bias dependence of the self-consistent channel potential of the two devices, and then use the result of that comparison to show how the difference in the densities of states in the two materials results in the variation in channel charge depicted in Fig. 2.7 and hence in the linearity according to (2.5).

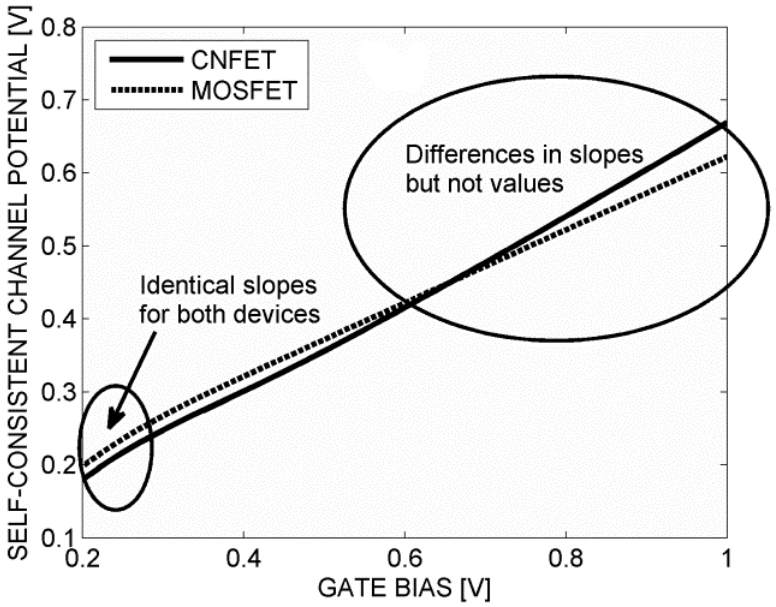
i) Bias Dependence of Self-Consistent Channel Potential

Fig. 2.8(a) shows the result for the self-consistent channel potential v_{SCF} as a function of the gate bias voltage v_G; as shown, the values of v_{SCF} at each v_G are approximately equal in the two devices. To understand this result, which is not entirely obvious, we first note that for typical structures and for sufficiently large drain bias, the large-signal input equivalent circuit looking into the gate consists primarily of the series combination of C_{ge} and C_{sq}, with the latter being a nonlinear function of the voltage v across it. The stored charge on each of these two capacitors must be equal and represents the channel charge λ:

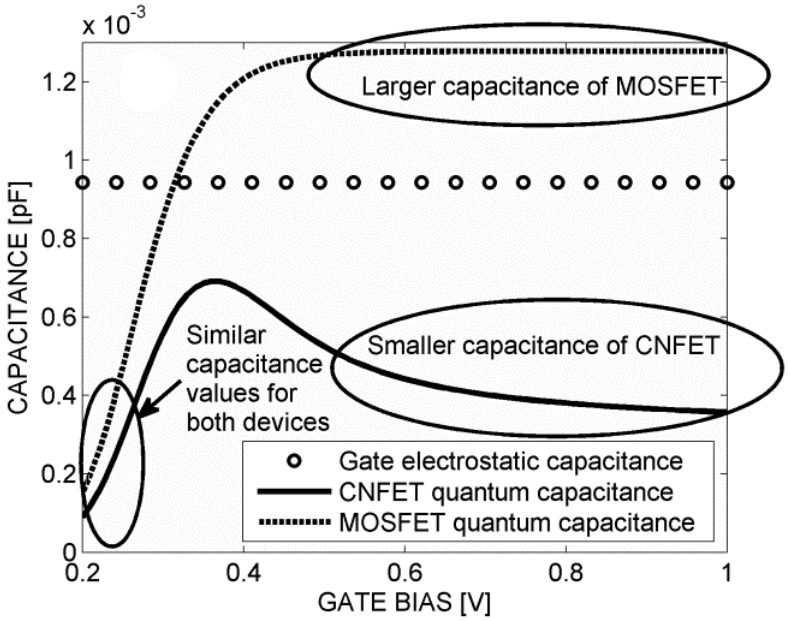
$$C_{\mathrm{ge}}(v_G - v_{\mathrm{SCF}}) = \lambda = \int_0^{v_{\mathrm{SCF}}} C_{\mathrm{sq}}(v)\, dv \tag{2.6}$$

where C_{sq} is written as $C_{\mathrm{sq}}(v)$ to emphasize its dependence on v. By differentiating both sides of (2.6) with respect to v_{SCF} and rearranging the result, we can find an expression for the rate of change of v_{SCF} with respect to v_G:

$$\frac{dv_{\mathrm{SCF}}}{dv_G} = \frac{C_{\mathrm{ge}}}{\left[C_{\mathrm{ge}} + C_{\mathrm{sq}}(v_{\mathrm{SCF}})\right]} \tag{2.7}$$



(a)



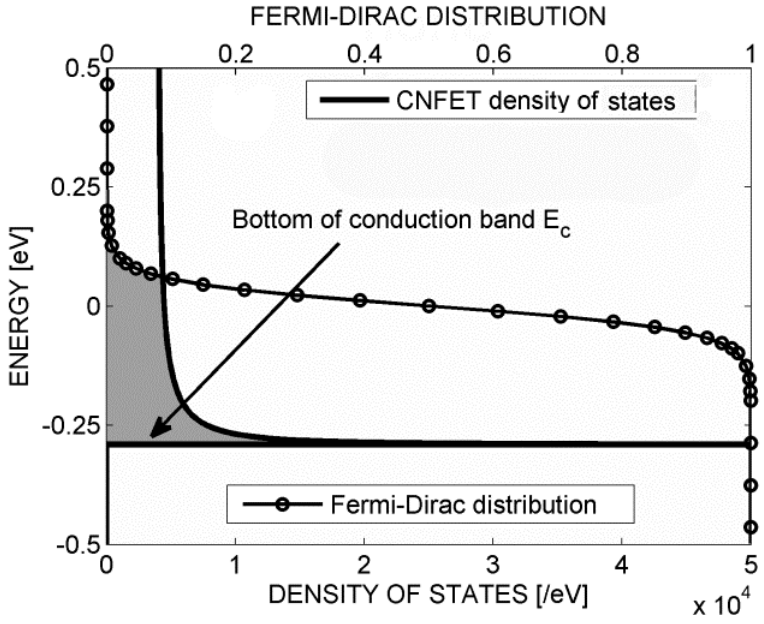
(b)

Fig. 2.8. (a) Self-consistent channel potential and (b) electrostatic and quantum capacitances (C_{ge} and C_{sq}) vs. gate bias for a 20-nm ballistic array-based CNFET and ballistic MOSFET.

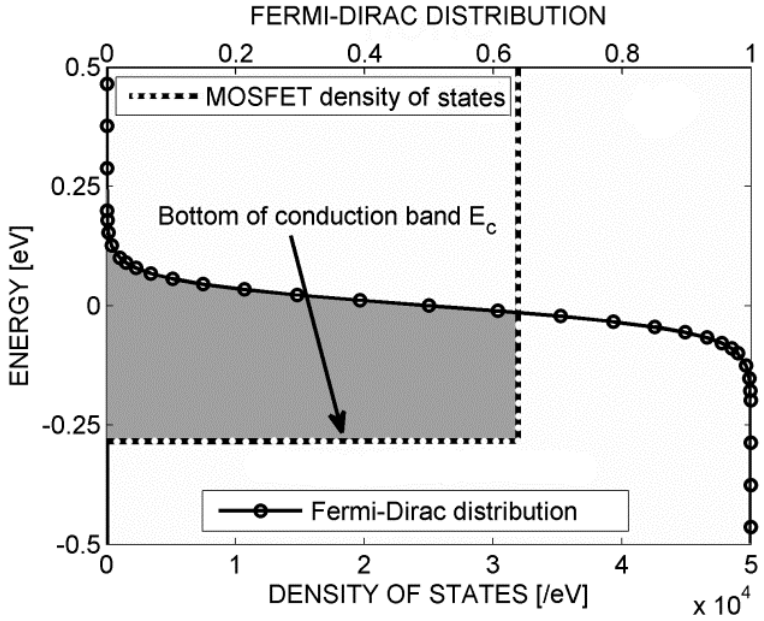
Fig. 2.8(b) shows C_{sq} and C_{ge} vs. gate bias for both the CNFET and the MOSFET. As shown, C_{ge} is the same constant value in the two devices. On the other hand, the behavior of C_{sq} depends on the density of states; once the gate bias is high enough [above 0.2 V for the devices considered in this work] to push the top of the barrier down to the source Fermi level, a plot of C_{sq} vs. gate bias will closely follow the shape of the density of states vs. energy. For the CNFET (governed by a 1-D density of states), we can thus expect C_{sq} to rise to a peak and then rapidly diminish, while for the MOSFET (governed by a 2-D density of states), we can expect C_{sq} to rise to a constant value. This behavior is indeed present in Fig. 2.8(b), but appreciable differences in the resulting values of C_{sq} are observable only for gate biases beyond ~ 0.35 V. According to (2.7), v_{SCF} will thus rise with v_G in a similar fashion in the two devices at least until $v_G \sim 0.35$ V; beyond this point, first-order differences in the *slopes* of the v_{SCF} vs. v_G curves do occur, consistent with (2.7), but they are insufficient to create appreciable differences in the *values* of v_{SCF} at each v_G.

ii) Effect of Density of States on Channel Charge

The self-consistent potential v_{SCF} determines the position of the conduction-band edge at the top of the barrier with respect to the source Fermi level [66], and since we have established that v_{SCF} is approximately the same at each v_G in the two devices, the band edges will be aligned. The reason for lower charge in the CNFET at high bias then readily follows from Fig. 2.9, where the density of states for the two devices is plotted for $v_G = 0.8$ V, along with the source Fermi function; in the figures, the source Fermi level is at 0, and the conduction-band edges are just below -0.25 eV. The charge is given by the area under the product of the Fermi function and the density of states, as shown by the shaded regions of Fig. 2.9. The diminishing density of states above the band edge in the CNFET [Fig. 2.9(a)] causes the area (charge) to be much lower than that in the MOSFET [Fig. 2.9(b)], which has a constant density of states above the band edge.



(a)



(b)

Fig. 2.9. Density of states at the top of the barrier and the source Fermi function vs. energy, plotted for gate and drain bias voltages both equal to 0.8 V, at room temperature (300 K) for a 20-nm (a) ballistic array-based CNFET and (b) ballistic MOSFET. The source Fermi level is taken as the reference (set to 0) and the conduction-band edge E_C, which is common for the two devices, is marked. The shaded region indicates the overlap of the Fermi function with the density of states and its area indicates the approximate amount of charge in the channel.

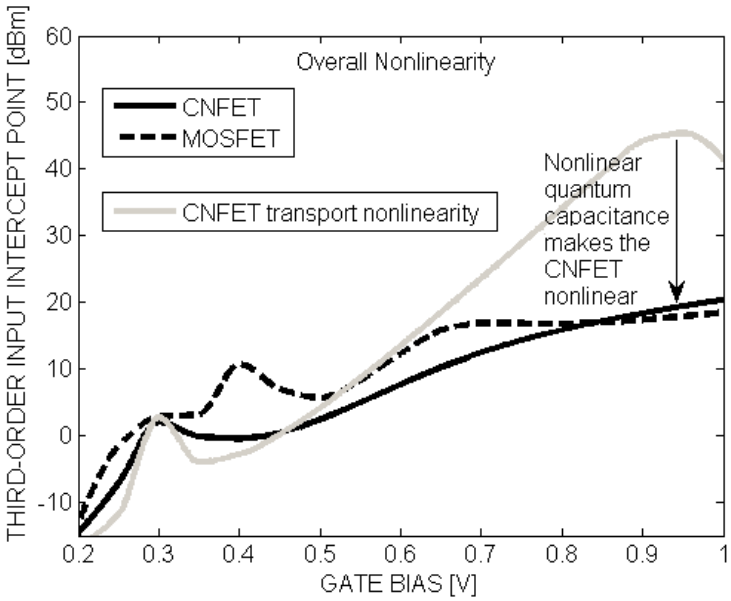


Fig. 2.10. Intrinsic IIP3 vs. gate bias for a 20-nm ballistic array-based CNFET and ballistic MOSFET.

Thus, despite being equivalent in all other aspects (*e.g.*, gate electrostatic capacitance, dimensions, equilibrium source-drain barrier height, and applied bias), the lower density of states in the CNFET ultimately causes it to have less channel charge at high gate bias, leading to (2.5) being better satisfied, and hence leading to better linearity. However, this discussion assumes that the transport current is the only source of nonlinearity in the device.

c) Overall Nonlinearity

So far, we have considered the nonlinearity due only to the transport current. The quantum capacitance is another major source of nonlinearity in these devices that must be considered. Fig. 2.10 shows the *overall* IIP3 of the two devices, *i.e.*, it shows the IIP3 found from MWO with all the coefficients in (2.1)-(2.4) retained. Most interestingly, with the inclusion of all the nonlinear elements, the linearity of the CNFET at high gate bias substantially degrades, and it becomes comparable to (or even worse than) that of the MOSFET. The reason behind this linearity degradation is the nonlinear behavior of the quantum capacitance C_{sq}. In the CNFET, the same 1-D density of states that causes (2.5)

to be well satisfied (as discussed in conjunction with Fig. 2.9) also yields a smaller and highly nonlinear C_{sq} in comparison to that of the MOSFET. Fig. 2.11 shows the quantum capacitances in the two devices as a function of the self-consistent channel potential v_{SCF}; the figure is very similar to Fig. 2.8(b), owing to the essentially linear dependence of v_{SCF} on v_G [Fig. 2.8(a)]. By definition, the value of capacitance C_{sq} at each point on the curve is the first coefficient C_{sq1} in the expression of the nonlinear small-signal charge q_{sq} in (2.1). The second and third coefficients, C_{sq2} and C_{sq3}, respectively, are therefore the first and second derivatives of the curves in Fig. 2.11. At high bias ($v_G \geq 0.5$ V, *i.e.*, $v_{\mathrm{SCF}} \geq 0.35$ V), the highly nonlinear curve for the CNFET results in pronounced values for C_{sq2} and C_{sq3}, and hence a highly nonlinear q_{sq} in (2.1), leading to substantial linearity degradation. On the other hand, the constant value of C_{sq} in the MOSFET at high bias, which is a direct outcome of the constant 2-D density of states for the MOSFET, results in essentially zero values for C_{sq2} and C_{sq3}, reducing (2.1) for q_{sq} into an essentially linear equation; in fact, the overall IIP3 of the MOSFET in Fig. 2.10 is not very different from the IIP3 due only to the transport current in Fig. 2.6.

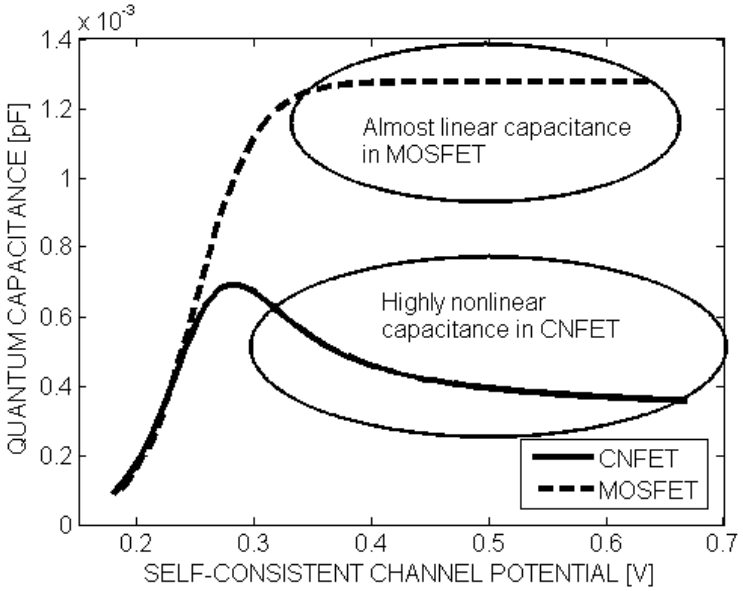
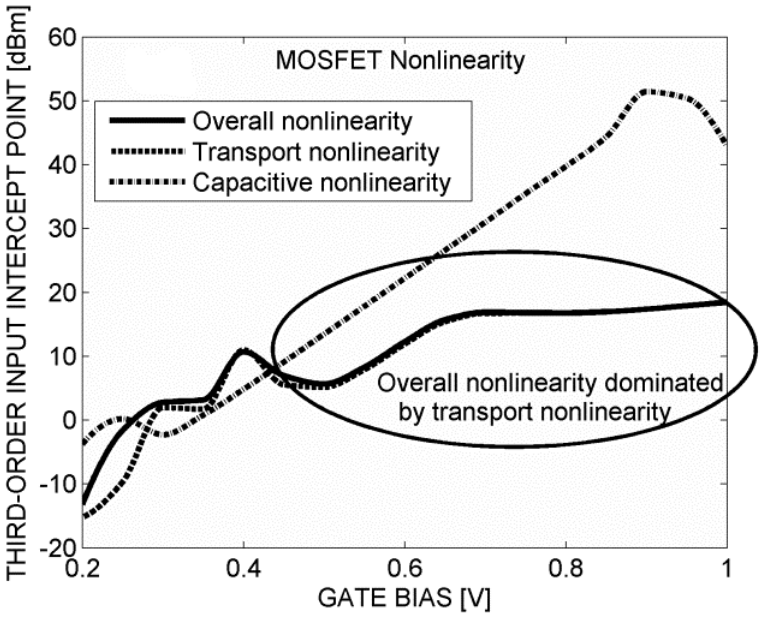
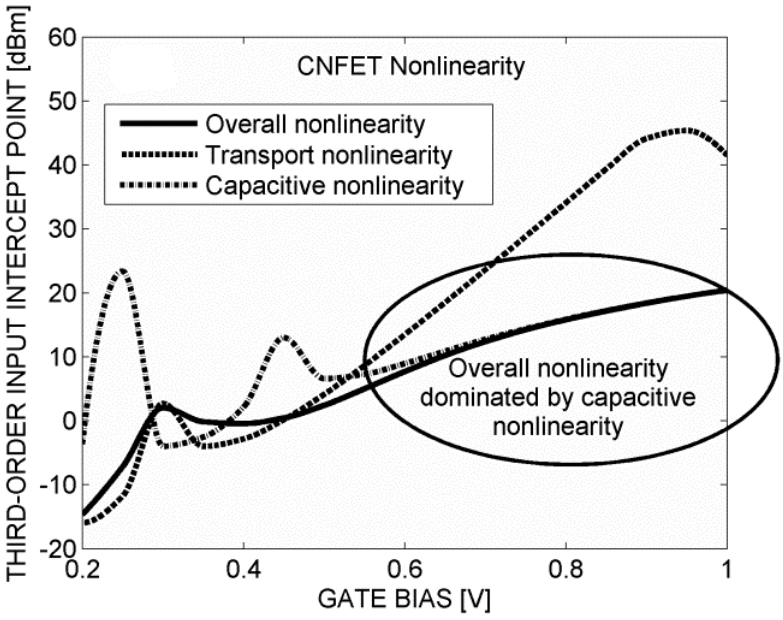


Fig. 2.11. Quantum capacitance C_{sq} vs. self-consistent channel potential for a 20-nm ballistic array-based CNFET and ballistic MOSFET.



(a)



(b)

Fig. 2.12. Intrinsic IIP3 vs. gate bias for (a) a 20-nm ballistic MOSFET and (b) a 20-nm ballistic array-based CNFET, with the different sources of nonlinearity selectively included.

Based on our study, a key difference between the sources of nonlinearity in a ballistic MOSFET and CNFET can be identified. We have summarized the results in Figs. 2.12(a) and (b), which show the overall IIP3 of each device along with the IIP3 found by selectively excluding either the transport-current or quantum-capacitance nonlinearities. Fig. 2.12(a) shows that the nonlinearity of the MOSFET remains relatively unaffected when the nonlinearity of the quantum capacitance is neglected, indicating that the transport current is the major source of nonlinearity in a MOSFET. On the other hand, Fig. 2.12(b) shows that the nonlinearity of the CNFET at high bias remains unaffected if the nonlinearity of the transport current is neglected, indicating that the nonlinear quantum capacitance is a major source of nonlinearity in a CNFET.

2.3.3. Linearity of Array-Based CNFETs

In this section, we focus on the extrinsic linearity of CNFETs, *i.e.*, the linearity found including the effects of the device parasitics. We examine the impact of various parameters of an array-based structure, including device width, gate oxide, and tube pitch, with the aim of understanding their potential in enhancing device linearity.

a) Calculation of Parasitics

To calculate the extrinsic parasitics, the gate metal is assumed to be made out of tungsten with dimensions $W \times L_g \times t_g$ of $1\ \mu\text{m} \times 20\ \text{nm} \times 50\ \text{nm}$; the $4.24\sim5.3$ eV [67] work function of tungsten is comparable to that of an intrinsic nanotube [68]. Based on the existing literature on resistivity of RF sputtered tungsten thin films [69], the total film resistance R_g of the gate metal is calculated to be $200\ \Omega$ at 24 GHz; the distributed gate resistance can then be modeled as a lumped resistance $R_{\text{g,eff}} = R_g/3$. The source/drain contacts are assumed to be made of yttrium with dimensions $W \times L_{\text{s/d}} \times t_{\text{s/d}}$ of $1\ \mu\text{m} \times 20\ \text{nm} \times 20\ \text{nm}$. Yttrium has been experimentally used to form an ohmic contact with the conduction band of a nanotube [54]. Based on the resistivity of yttrium thin films in [70], the total film resistances R_s and R_d for the source and drain contacts are calculated to be $1.77\ \Omega$.

Simulating the open-pad structure in COMSOL, the parasitic capacitances are calculated to be $C_{\mathrm{gs,ext}} = 20.7$ aF , $C_{\mathrm{gd,ext}} = 20.8$ aF , and $C_{\mathrm{sd,ext}} = 4.99$ aF. Table 2.1 lists the intrinsic and extrinsic component values of the 100-tube, 1-μm wide CNFET studied in this work; the bias-dependent values are shown for gate and drain biases both equal to 0.8 V.

TABLE 2.1. INTRINSIC AND EXTRINSIC CIRCUIT COMPONENTS OF THE ARRAY-BASED CNFET

Intrinsic Components							**Extrinsic Components**				
C_{ge} **[aF]**	C_{se} **[aF]**	C_{de} **[aF]**	C_{sq1} **[aF]**	C_{dq1} **[aF]**	g_{sq1} **[mA/V]**	g_{dq1} **[mA/V]**	$C_{\mathrm{gs,ext}}$ **[aF]**	$C_{\mathrm{gd,ext}}$ **[aF]**	$C_{\mathrm{sd,ext}}$ **[aF]**	$R_{\mathrm{g,eff}}$ **[Ω]**	R_s, R_d **[Ω]**
900	120	1	380	4.09×10^{-6}	15.5	4.23×10^{-8}	20.7	20.8	4.99	66.67	1.77

b) Effect of Gate Oxide

Although the discussion in Section 2.3.2 above clearly indicates that it is the nonlinear quantum capacitance that limits CNFET linearity, it is nevertheless interesting to test the outcome of satisfying condition (2.5) in a realistic extrinsic structure, which should at least linearize the transport current.

To do so, we consider an extrinsic array-based structure with the gate oxide changed from hafnium oxide to a fictitious oxide that has a tenfold increase in relative permittivity. This increase in the relative permittivity will aid in achieving the derived condition for "inherent linearity" in (2.5) by increasing C_{ge}. Fig. 2.13 shows the linearity figures of merit for both the original ($\varepsilon_r = 16$) and new ($\varepsilon_r = 160$) transistors, obtained from MWO and the full extrinsic circuit of Fig. 2.5.

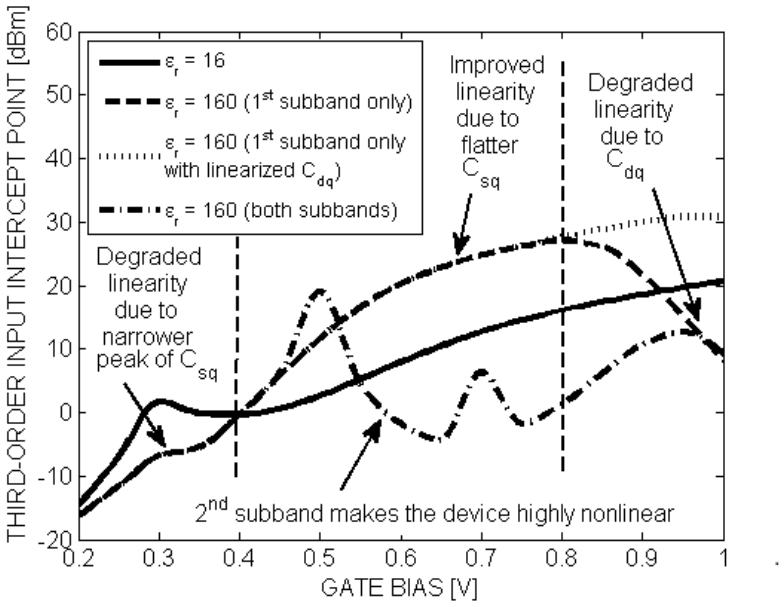


Fig. 2.13. Extrinsic IIP3 vs. gate bias for a 20-nm ballistic array-based CNFET with different gate oxides.

The results in Fig. 2.13 are best understood by considering the three bias regions delineated in the figure and focusing on the solid ($\varepsilon_r = 16$) and dashed ($\varepsilon_r = 160$, 1st subband only) curves, leaving aside the dotted and stippled curves for the moment. The following observations can be made.

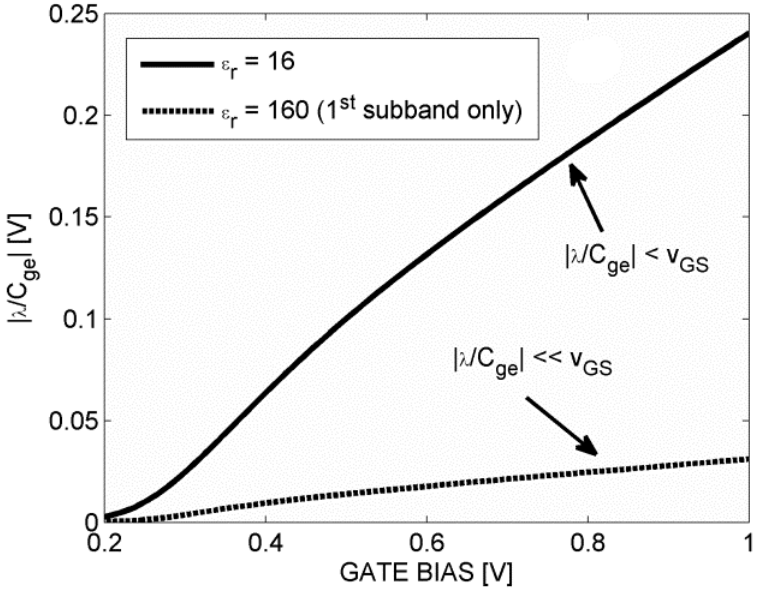
- In the central-bias region ($0.4\,\text{V} \leq v_G \leq 0.8\,\text{V}$), the CNFET does become more linear, and this is a result of condition (2.5) being better satisfied in the presence of a larger oxide capacitance, as illustrated by the curves in Fig. 2.14(a). However, in the low-bias ($v_G \leq 0.4\,\text{V}$) and high-bias ($v_G \geq 0.8\,\text{V}$) regions, the linearity degrades.

- The degradation at low bias can be explained by the behavior of the source quantum capacitance C_{sq}, which is also the first coefficient in the expression of q_{sq} in (2.1) and which is plotted in Fig. 2.14(b) vs. gate bias. In both cases ($\varepsilon_r = 16$ and $\varepsilon_r = 160$), the curves in Fig. 2.14(b) rise sharply and then gradually fall as the gate bias is increased, which is consistent with the shape of the 1-D density of

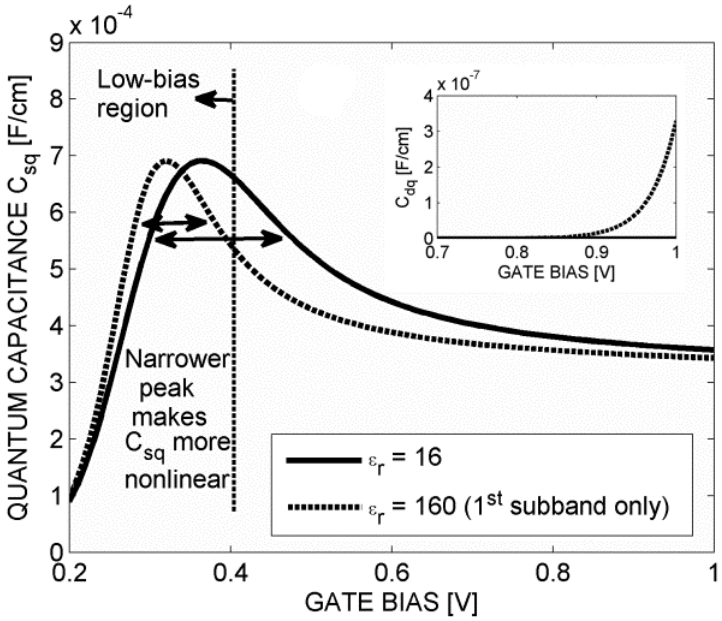
states of the nanotube. However, in the case of the larger ε_r, an applied gate bias has a superior control over the channel potential, *i.e.*, the conduction-band edge is more easily pushed down by the gate bias, causing the C_{sq} curve for the CNFET to have a *squeezed* shape (narrower peak) as the conduction-band states get filled faster. The squeezed shape makes the curve more nonlinear, leading to substantial linearity degradation at low bias ($v_G \leq 0.4$ V).

- The degradation at large gate bias ($v_G \geq 0.8$ V) can be explained by the influence of the *drain* quantum capacitance. In this case, when ε_r is large, the applied gate bias can push the conduction band in the channel low enough to cause a significant amount of *drain*-injected electrons to enter the channel. As a result, the *drain* quantum capacitance C_{dq} acts as a second source of capacitive nonlinearity in the device (over and above C_{sq}), reducing the IIP3. The effect is pronounced only at sufficiently large gate bias ($v_G \geq 0.8$ V), where C_{dq} becomes appreciable, as shown in the inset to Fig. 2.14(b); the corresponding impact on the IIP3 is emphasized by the dotted curve in Fig. 2.13, where we have plotted the IIP3 at high bias in the presence of a *linearized* C_{dq}.

Overall, the results in Fig. 2.13 demonstrate that an increase in gate-oxide capacitance C_{ge} does have the potential to improve linearity, subject to the caveat of degradations at low and high bias arising from the impacts of the quantum capacitances, C_{sq} and C_{dq}.



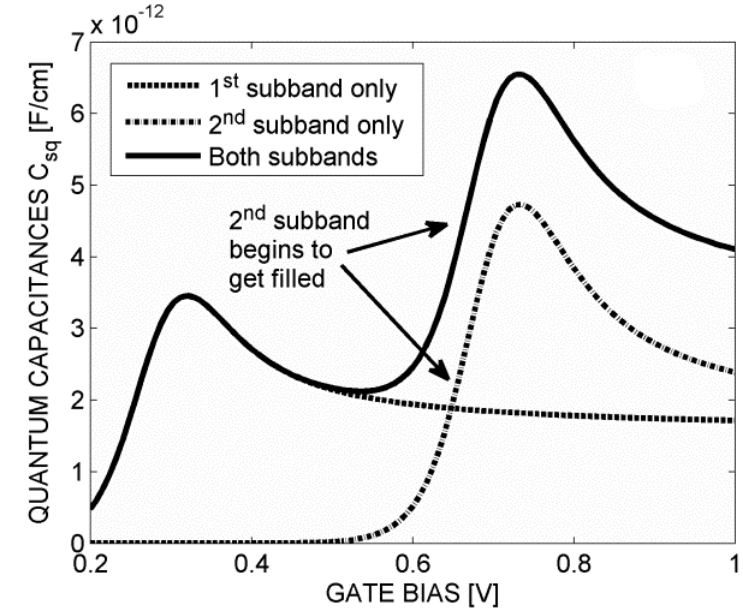
(a)



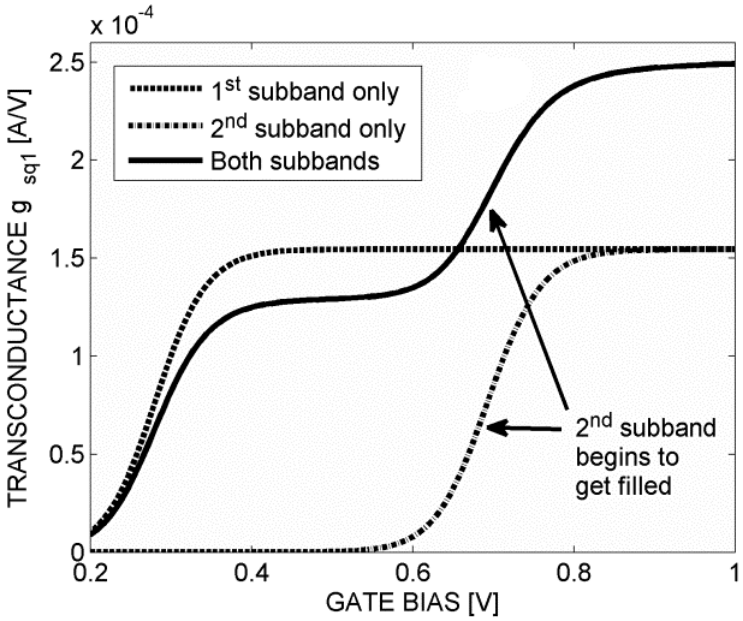
(b)

Fig. 2.14. (a) The ratio $|\lambda/C_{ge}|$ related to condition (5) for "inherent linearity" and (b) plot of source quantum capacitance C_{sq} vs. gate bias for a 20-nm ballistic array-based CNFET with different gate-oxide values. The inset of (b) shows the drain quantum capacitance C_{dq} vs. gate bias.

c) Impact of the Second Subband



(a)



(b)

Fig. 2.15. (a) Source quantum capacitance C_{sq} and (b) transconductance g_{sq1} vs. gate bias for an array-based CNFET with large gate electrostatic capacitance C_{ge}.

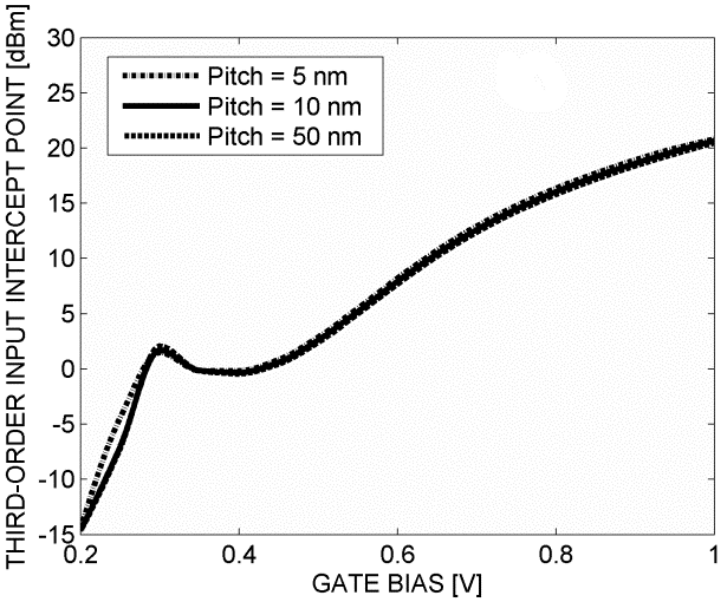
So far in our simulations, we have neglected the second subband because it does not come into play for the CNFET structure considered in this work, *i.e.*, with HfO_2 as the gate oxide ($\varepsilon_r = 16$), until very high gate bias ($v_G > 1.0$ V). However, it should be noted that for the case of the very large ε_r just discussed, the second subband does come into play much earlier and needs to be considered in order to capture the overall linearity of the CNFET. The stippled curve in Fig. 2.13 shows that when the second subband is taken into account, the IIP3 of the CNFET with higher gate capacitance ($\varepsilon_r = 160$) goes *below* the IIP3 of the CNFET with hafnium oxide ($\varepsilon_r = 16$) for gate biases in excess of 0.5 V. This behavior can be explained by the plots of the source quantum capacitance C_{sq} and the transconductance g_{sq1} appearing in (2.3), as shown in the two parts of Fig. 2.15; as illustrated, when the second subband begins to get filled, C_{sq} and the g_{sq1} become highly nonlinear, and this can cause the overall nonlinearity to be *worse* with a high oxide capacitance. The potential impact of the second subband should be kept in mind if an increase in C_{ge} is used in an attempt to improve the linearity according to (2.5).

d) Effect of Tube Pitch and Channel Width

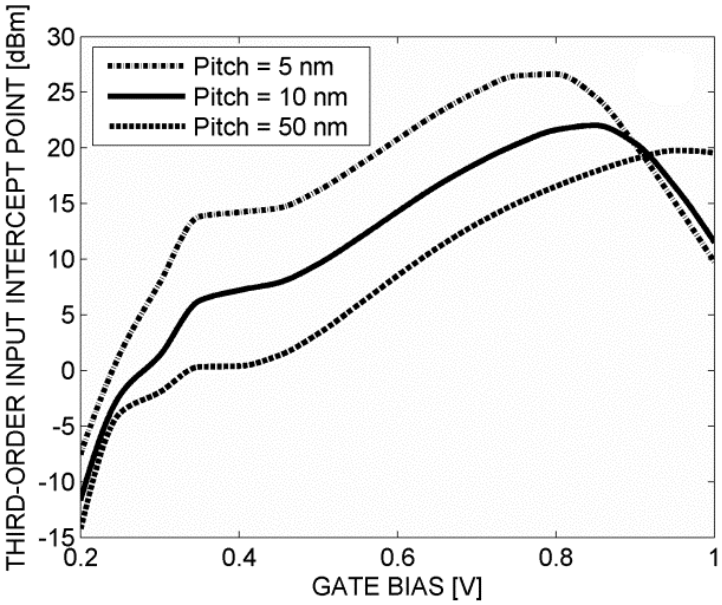
So far in this work, we have assumed a channel width of 1µm and a tube pitch of 10 nm. Generally, a low tube pitch is preferred in array-based structures, since it has been reported to offer multiple benefits, ranging from improved cutoff frequencies to a larger drive current and better impedance matching for the overall CNFET [18, 71]. It is hence important to assess the impact of tube pitch on the RF linearity.

Fig. 2.16 (a) shows the IIP3 vs. gate bias for three 1-µm wide array-based CNFETs with varying tube pitch. It is evident from the plots that the linearity of the device is almost insensitive to the tube pitch. However, for devices with wider channels (above 30 µm), where the extrinsic components are more pronounced than in the 1-µm wide device, the linearity improves considerably with a reduction in the tube pitch [Fig. 2.16 (b)]. This improvement can be attributed to the linearizing feedback [p.101, 43] provided by the extrinsic circuit components, an effect that becomes more pronounced as the number of tubes increases (*i.e.*, as the pitch decreases) and as the drive current through the

extrinsic components thus increases; we found the extrinsic components $R_{\mathrm{g,eff}}$ and $C_{\mathrm{gd,ext}}$ to be primarily responsible for the improvement.



(a)



(b)

Fig. 2.16. Extrinsic IIP3 vs. gate bias of (a) 1-µm wide and (b) 30-µm wide array-based CNFETs for different tube pitches.

A low tube pitch, which has already been flagged in the literature [18, 71] as being important for a variety of device metrics, can thus also be flagged as aiding device linearity (when the device width is sufficiently large). It is worth mentioning that although we would expect similar behavior in a MOSFET with respect to channel width (*i.e.*, better linearity arising from the larger extrinsic parasitics of a wider device), the linearity enhancement due to tube pitch is unique to the CNFET.

2.4. Conclusions

The following conclusions can be drawn regarding the linearity of an array-based CNFET [Fig. 2.1(a)] vs. its MOSFET counterpart [Fig 2.1(c)]:

1. Under the assumption of ballistic transport, the linearity of an array-based CNFET is comparable to that of its MOSFET counterpart (having similar dimensions and electrostatic capacitances), as shown by the results in Fig. 2.10.
2. The highly nonlinear quantum capacitance is the main source of nonlinearity in a CNFET, whereas the transport nonlinearity dominates the linearity of its conventional MOSFET counterpart (Fig. 2.12).
3. The array-based CNFET does not become "inherently linear" even under extreme situations, such as the case of a huge oxide capacitance [22], as shown in Fig. 2.13; extenuating factors include the impact of the source and drain quantum capacitances and the impact of the second subband.
4. The tube pitch of the array-based CNFET does not affect the linearity for narrow devices ($W \ll 30$ μm), as shown in Fig. 2.16 (a). However, for wider devices ($W \geq 30$ μm), the CNFET linearity increases as the tube pitch is decreased [Fig. 2.16 (b)]. This is a rather favorable observation since it implies that lowering the tube pitch to increase the drive current and improve the cutoff frequency will also improve the linearity.

While CNFETs may not be "inherently linear," they should offer at least comparable linearity to their MOSFET counterparts. Given the already reported potential for superior

f_T and $f_{\max}$ of array-based CNFETs, they hence continue to be promising candidates for RF applications.

Chapter 3

A Modified Top-of-the-Barrier Solver for Graphene and Its Application to Predict RF Linearity[8]

3.1. Introduction

The excellent electronic properties of graphene are believed to make it a promising alternative to silicon for use in future electronics, particularly for analog circuit applications. As the down-scaling of graphene channels continues, compact modeling approaches that can tractably predict the terminal behavior, including effects arising from the zero bandgap --- such as variations in the densities of states between the channel and source and drain regions [66] and band-to-band tunneling [29] --- will be essential to explore graphene's circuit capabilities.

To date, the amount of work done on modeling GFETs has been significant. The reported approaches range from semi-classical, top-of-the-barrier models (TBMs) [13, 27] to numerically involved quantum-mechanical solvers based on non-equilibrium Green's functions (NEGF) [10, 72-74]. The quantum-mechanical solvers, although complex and numerically demanding, are necessary to provide rigorous theoretical

[8] A version of this chapter has been published [16].

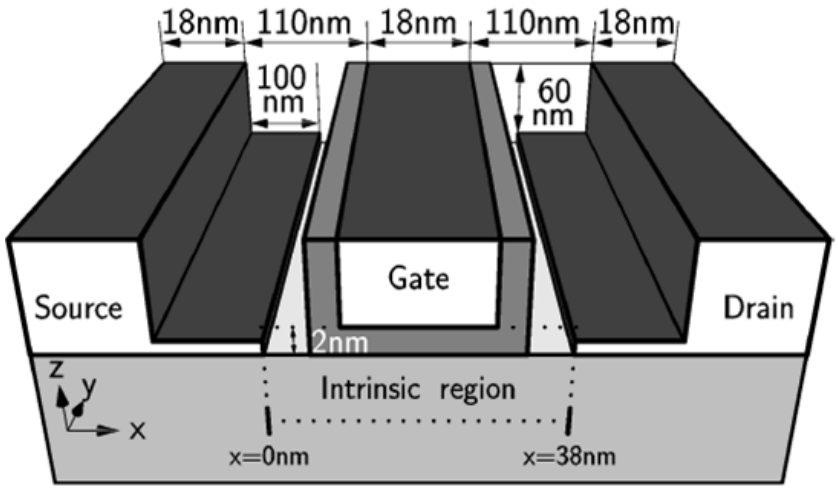
benchmarks, while semi-classical TBMs can provide reasonably accurate results when speed and simplicity matter, such as for compact models.

In this work, we present a way to simulate electronic transport in GFETs through a *modified* TBM [13] that captures both variations in the densities of states and band-to-band tunneling, while remaining numerically efficient. The new model is shown to produce accurate results when compared to a more rigorous, self-consistent, quantum-transport solver based on NEGF [10], and its potential is demonstrated by investigating the RF linearity of GFETs. RF linearity is an important transistor property that is relevant for a variety of circuit applications, but which is notoriously difficult to predict, requiring both accuracy and tractability in the modeling approach.

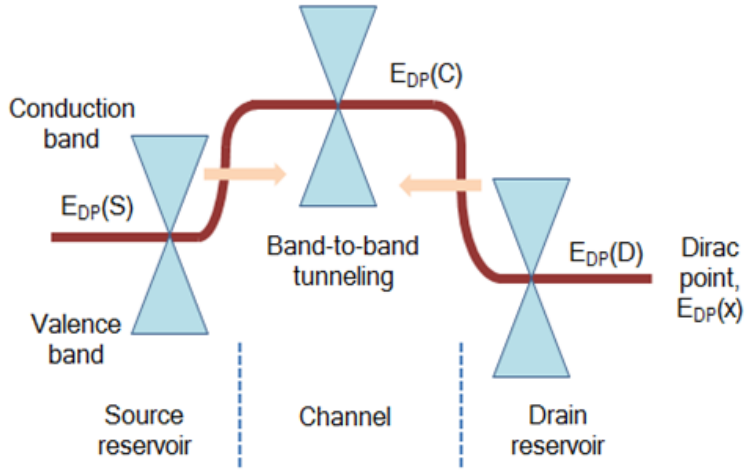
Section 3.2 explains the implementation of the short-channel, quantum-mechanical, and heavily doped source/drain effects within the top-of-the-barrier method and implementation of the corresponding nonlinear small-signal circuit. Section 3.3 shows the accuracy of the amended TBM and compares the linearity of GFETs with their MOSFET and CNFET counterparts of similar specifications; the effect of biasing on linearity is also discussed. The conclusions of this work are summarized in Section 3.4.

3.2. Theory

As graphene devices mature beyond the research stage, doped MOSFET-style devices will be necessary to obtain better performance than the Schottky-barrier devices prevalent today [35, 75]. Based on this observation, and as a starting point, we assume the device geometry shown in Fig. 3.1(a), consisting of an intrinsic channel region surrounded by heavily doped source and drain regions. For reference throughout this discussion, a plot of the Dirac-point energy E_{DP} versus position x under typical operating conditions in such a device is shown in Fig. 3.1(b).



(a)



(b)

Fig. 3.1. (a) Device geometry of the simulated GFET [10]. (b) Dirac-point energy $E_{DP}(x)$ versus position x, where the superimposed cones represent graphene's band structure.

Top-of-the-barrier approaches [13] have been widely used to model ballistic transport in field-effect transistors. Despite its simplicity, the conventional TBM has been shown to offer good accuracy when compared to more rigorous simulation methodologies for CNFETs and MOSFETs [13, 27]. The most striking difference between the channel materials used in such FETs and graphene is that the bandgap is zero in graphene. The zero bandgap causes the transport to be strongly ambipolar, *i.e.*, both electrons and holes

contribute to the charge flow. This ambipolar character of graphene can be taken into account by modifying the transport equations of the conventional TBM [13] in the following manner:

$$n_{\mathrm{S/D}} = \frac{1}{2}\int_{E_{\mathrm{DP}}(C)}^{\infty} D_C(E) f(E - E_{\mathrm{F,S/D}})\mathrm{d}E \tag{3.1}$$

$$p_{\mathrm{S/D}} = \frac{1}{2}\int_{-\infty}^{E_{\mathrm{DP}}(C)} D_C(E)[1 - f(E - E_{\mathrm{F,S/D}})]\mathrm{d}E \tag{3.2}$$

$$Q = q(p_S + p_D) - q(n_S + n_D) \tag{3.3}$$

$$I_{\mathrm{S/D}} = q v_F (p_{\mathrm{S/D}} - n_{\mathrm{S/D}}) \tag{3.4}$$

$$I_{\mathrm{DS}} = I_D - I_S \tag{3.5}$$

where $n_{\mathrm{S/D}}$ and $p_{\mathrm{S/D}}$ are the electron and hole density of the channel arising from contributions from the source/drain (S/D), respectively; $E_{\mathrm{DP}}(C)$ is the channel Dirac-point energy, determined by shifting its equilibrium position $E_{\mathrm{DP,E}}(C)$ by the self-consistent potential U_{SCF}, *i.e.*, $E_{\mathrm{DP}}(C) \equiv E_{\mathrm{DP,E}}(C) - U_{\mathrm{SCF}}$; $D_C(E)$ is the channel density of states (DOS); $f(E)$ is the Fermi-Dirac distribution function; $E_{\mathrm{F,S/D}}$ is the chemical potential of the source/drain reservoir; Q is the total charge in the channel, which is used in Poisson's equation to determine U_{SCF}; $I_{\mathrm{S/D}}$ is the conventional current from the source/drain; and I_{DS} is the net drain-to-source current.

Although the TBM specified by (3.1)-(3.5) can capture the ambipolar nature of transport in graphene, it still misses important physics. First, the lack of a bandgap in graphene results in significant band-to-band tunneling, a quantum-mechanical effect that the conventional TBM does not include. Second, the conventional TBM cannot model the difference in the DOS between the doped reservoirs and the channel. These limitations can be overcome by amending the TBM through the following steps. First, band-to-band tunneling is introduced to the TBM with the WKB approximation [29]. Second, we construct an "effective DOS" that accounts for the fact that the densities of

states in the source and drain reservoirs vanish at the local Dirac points [$E_{\mathrm{DP}}(S)$ and $E_{\mathrm{DP}}(D)$ in Fig. 3.1(b)], preventing any transmission at these energies [10].

3.2.1. Effective Density of States

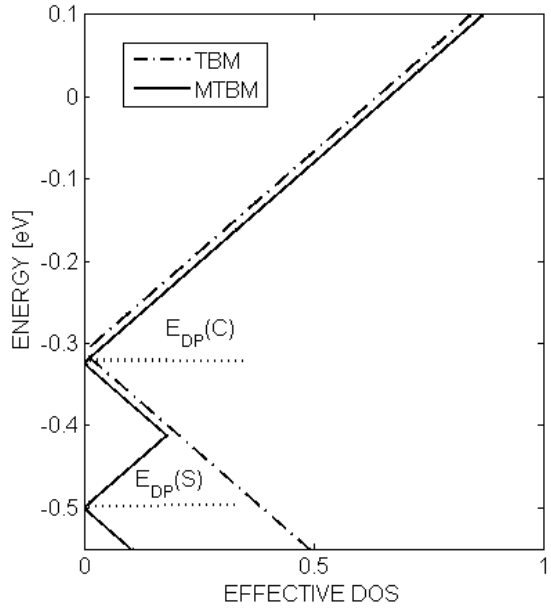
The effect on transport due to the lack of states at the Dirac points in the source and drain reservoirs can be included in the top-of-the-barrier formalism by replacing the *channel* density of states $D_C(E)$ with two *effective* densities of states. Our refined model hence replaces $D_C(E)$ with $D_{\mathrm{SC}}(E)$ for transport from the source to the channel and $D_{\mathrm{DC}}(E)$ for transport from the drain to the channel. At each energy, the effective DOS functions contain the minimum density of states existing between the corresponding reservoir (source or drain) and the channel. Equations (3.1) and (3.2) can therefore be re-written as

$$n_{\mathrm{S/D}} = \frac{1}{2}\int_{E_{\mathrm{DP}}(C)}^{\infty} D_{\mathrm{SC/DC}}(E) f(E - E_{\mathrm{F,S/D}})\mathrm{d}E \tag{3.6}$$

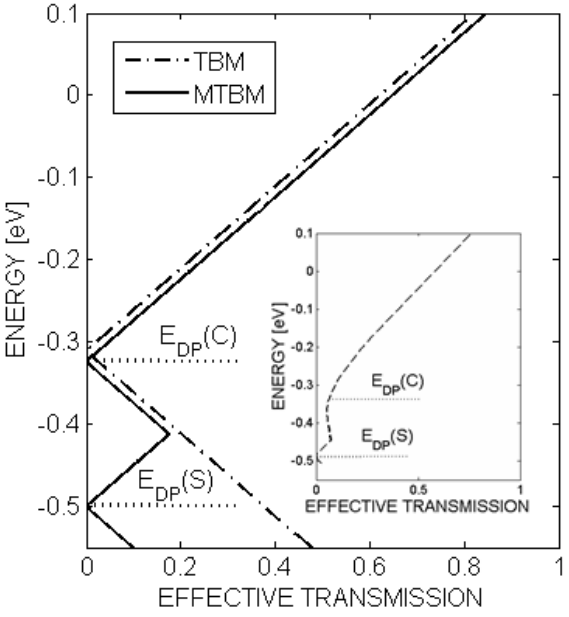
$$p_{\mathrm{S/D}} = \frac{1}{2}\int_{-\infty}^{E_{\mathrm{DP}}(C)} D_{\mathrm{SC/DC}}(E)[1 - f(E - E_{\mathrm{F,S/D}})]\mathrm{d}E \tag{3.7}$$

It is worth mentioning here that in the conventional TBM [13], the doped source and drain are assumed to be *infinite* reservoirs, in which case the "effective DOS" $D_{\mathrm{SC/DC}}(E)$ always reduces to the channel DOS $D_C(E)$, as it is the channel that determines (limits) the available states for transport at all energies.

Fig. 3.2(a) shows the effective DOS $D_{\mathrm{SC}}(E)$, which is zero at the Dirac-point energies associated with the source and the channel, and Fig. 3.2(b) shows the corresponding Landauer transmission function for the electron/hole transport, which is proportional to $D_{\mathrm{SC}}(E)$ for the energies shown; as illustrated, the transmission vanishes at those energies where $D_{\mathrm{SC}}(E)$ vanishes, in agreement with results obtained from the quantum-mechanical NEGF solver reported in [10]. For reference, the DOS and the transmission function from the *conventional* TBM are also plotted.



(a)



(b)

Fig. 3.2 (a) Effective (normalized) source-to-channel DOS $D_{SC}(E)$ of the *modified* TBM, being zero at the indicated Dirac-point energies, and (b) respective effective (normalized) transmission function describing the source-to-channel transport. (Inset: transmission obtained from a NEGF solver under similar operating conditions [10].) The DOS and the transmission function for the *conventional* TBM are also plotted for comparison.

3.2.2. Band-to-Band Tunneling

Graphene is a zero-bandgap semiconductor. This unique property makes it considerably different from its competitors, such as silicon and carbon nanotubes. The lack of a bandgap means that the valence band in the channel actively participates in transport along with the conduction band. For example, under certain gate biases, the carriers in the conduction band in the doped source and drain reservoirs can tunnel through the potential barrier at the corresponding reservoir-to-channel interface and enter the channel valence band [Fig. 3.1(b)]. This additional transport path, *i.e.*, in addition to regular thermionic transport over the barrier, must be taken into account in graphene to predict the terminal behavior.

In order to properly model the band-to-band tunneling in a GFET, we have modified the conventional TBM equations by introducing a tunneling probability $T_{\mathrm{KZ,S/D}}(E)$ as a function of energy E, so that (3.6) and (3.7) become

$$n_{\mathrm{S/D}} = \frac{1}{2}\int_{E_{\mathrm{DP}}(C)}^{\infty} T_{\mathrm{KZ,S/D}}(E) D_{\mathrm{SC/DC}}(E) f(E - E_{\mathrm{F,S/D}})\mathrm{d}E \tag{3.8}$$

$$p_{\mathrm{S/D}} = \frac{1}{2}\int_{-\infty}^{E_{\mathrm{DP}}(C)} T_{\mathrm{KZ,S/D}}(E) D_{\mathrm{SC/DC}}(E)\left[1 - f\left(E - E_{\mathrm{F,S/D}}\right)\right]\mathrm{d}E \tag{3.9}$$

where $T_{\mathrm{KZ,S/D}}(E)$ is the band-to-band (Klein-Zener) tunneling probability from the source/drain reservoir to channel. The Klein-Zener tunneling probability, unlike conventional quantum tunneling, can assume a value of unity when a carrier encounters a potential barrier, depending on its direction of travel, and within the WKB approximation, this tunneling probability $T_{\mathrm{KZ,S/D}}(k_{\perp})$ can be expressed as a function of the longitudinal component of the wave vector $k_{\perp}$ [29]:

$$T_{\mathrm{KZ,S/D}}(k_{\perp}) = \exp(-\pi {k_{\perp}}^2 \hbar v_F / q\varepsilon_{\mathrm{S/D}}) \tag{3.10}$$

where v_F is the Fermi velocity in graphene (~10^5 m/s), q is magnitude of the electronic charge, and $\varepsilon_{\mathrm{S/D}}$ is the electric field of the potential barrier at the source/drain reservoir-to-channel interface, given by $\varepsilon_{\mathrm{S/D}} = V_{\mathrm{diff,S/D}}/\Delta x$, with $V_{\mathrm{diff,S/D}}$ being the voltage across

the source/drain barrier and Δx being an effective width. By a change of variables, the tunneling probability can be rewritten as a function of energy:

$$T_{\mathrm{KZ,S/D}}(E) = \exp(-a/2)J_0(a/2) \tag{3.11}$$

where J_0 is the zeroth order Bessel function of the first kind and $a = \pi E^2/q\hbar^2 v_F \varepsilon_{\mathrm{S/D}}$.

These amendments lead to a modified set of TBM equations by which we are able to quantitatively match a GFET's behavior as predicted by more rigorous NEGF simulations [10], including the transistor's current-voltage curves, unity current-gain frequency f_T, transconductance g_m, and output conductance g_d. The usual fitting parameters [13] accompanied with the TBM, *i.e.*, α_g, α_d, and α_s, along with the new fit parameter Δx, are used to calibrate our model. The resulting solver is then used to determine the RF linearity of graphene using the method described in [15].

3.3. Results and Discussion

Our modeled GFET [Fig. 3.1(a)] has dimensions identical to the device investigated in [10], including a gate length of 18 nm. As a first step, the three fitting parameters α_g, α_d, and α_s were used to fit the TBM output with the NEGF data following the procedure of [13]. The fourth fitting parameter Δx was then introduced (accompanied by a tweaking of the initial values of α_g, α_d, and α_s) to further enhance the match. For the modeled GFET, we obtained $\alpha_g = 0.76$, $\alpha_d = 0.17$, $\alpha_s = 0.07$, and $\Delta x = 0.43$ nm.

Of particular note is the magnitude of α_d obtained during the fitting. The value in the conventional TBM is 0.033, a value much smaller than the value of 0.17 observed with our modified approach. The new value is consistent with results from [10], which predict a more dominant drain capacitance than is normally observed. The modified TBM thus more closely matches the underlying physics observed in much more detailed calculations, while still running in a fraction of the time.

It should also be noted that during fitting of both the conventional TBM and the modified TBM with NEGF data, an effective oxide thickness of 1.5 nm was needed for proper matching. This value differs from the actual value of 2 nm [indicated in Fig.

3.1(a)] to account for fringe capacitance in the real structure, which is automatically included in the NEGF but excluded in the TBM.

Fig. 3.3 shows the current-voltage characteristics of the GFET from our modified TBM solver and the self-consistent NEGF solver [10]. Considering the simplicity of our modified TBM, the agreement is quite remarkable. Fig. 3.4 demonstrates that the developed model not only matches the $i_D - v_G$ output characteristics, but also is in good agreement with more sensitive quantities involving derivatives, such as the transconductance g_m and output conductance g_d; in addition, our modified TBM shows a clear improvement towards the NEGF solution over the conventional TBM, particularly at lower v_G, where the magnitude of the tunneling current through the barrier is significant. Finally, we have plotted the second derivatives of the current in Fig. 3.5, which shows an even better improvement over the conventional model. Such improved agreement is important for an accurate prediction of the device's properties for analog circuit applications, such as RF linearity, which requires a proper modeling of the slope and curvature of the transistor characteristics versus voltage.

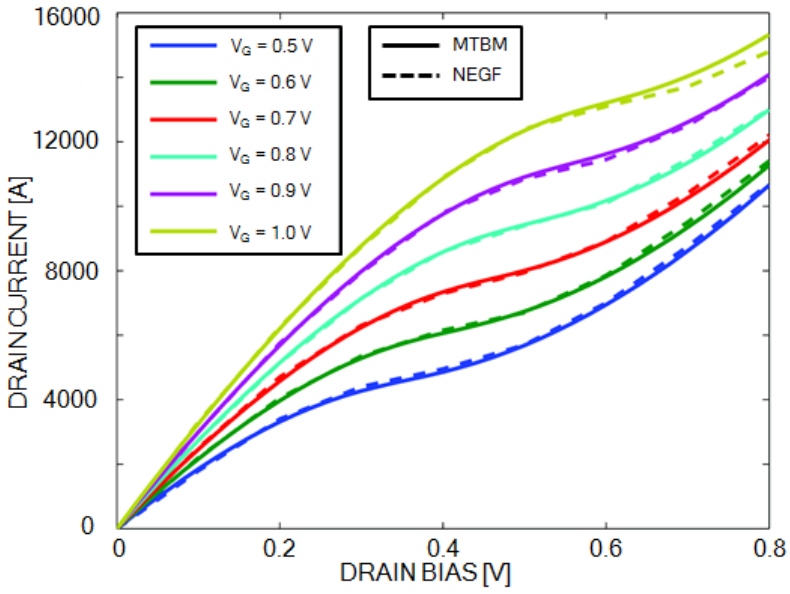
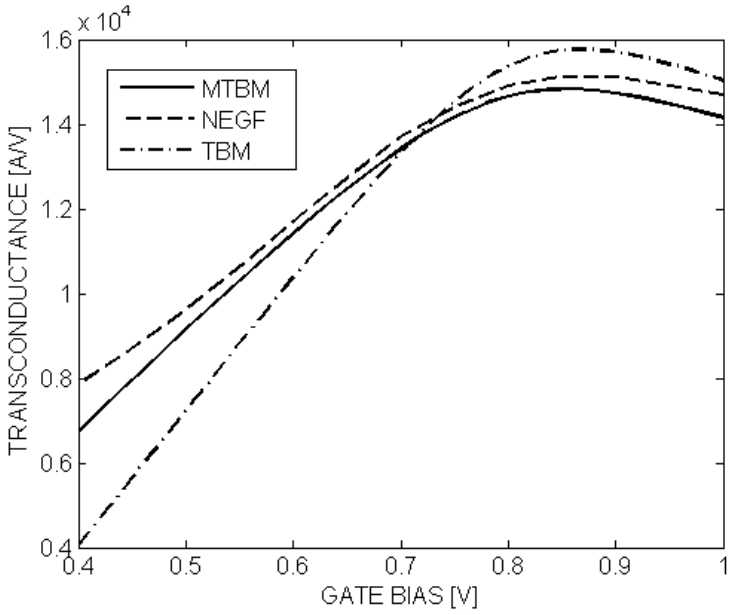
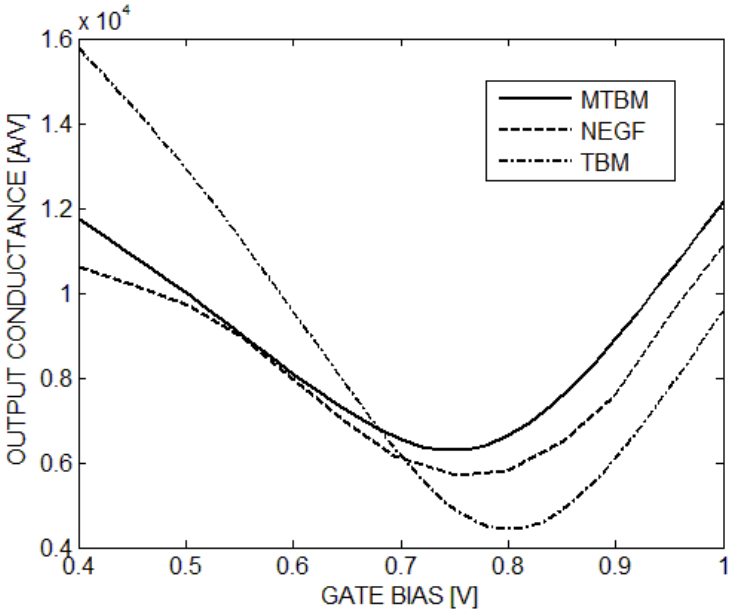


Fig. 3.3. Current-voltage characteristics of the GFET from NEGF and the modified top-of-the-barrier solver.



(a)



(b)

Fig. 3.4. (a) Transconductance and (b) output conductance versus gate bias of the GFET from NEGF, conventional top-of-the-barrier model, and modified top-of-the-barrier model.

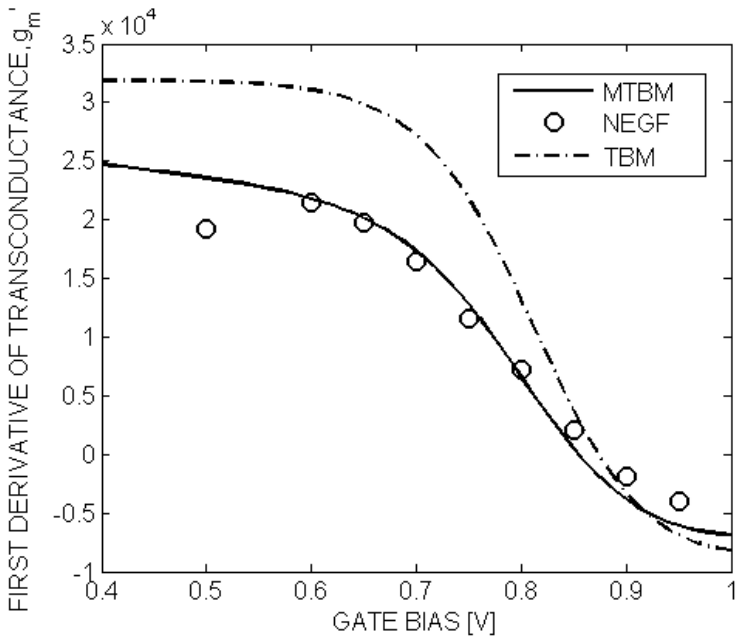


Fig. 3.5. First derivative of transconductance with respect to gate voltage of the GFET plotted versus gate bias. Shown are results from the conventional top-of-the-barrier model and modified top-of-the-barrier model. The modified TBM shows excellent agreement with NEGF data.

We used our modified TBM solver to determine the element values of a nonlinear, small-signal equivalent circuit (Fig. 3.6) for the GFET of Fig. 3.1(a), following the approach described in [15]. Here, C_{ge}, C_{se}, and C_{de} are the electrostatic capacitances of the GFET, which we assume to be linear as a first approximation, as discussed in [15]; C_{sq} and C_{dq} are the nonlinear source and drain quantum capacitances, respectively; and i_{ts} and i_{td} are the nonlinear current sources that model the quasi-static transport currents of the device. It is worth noting that due to the higher-order derivatives involved, accurately predicting these nonlinear elements requires the use of our modified TBM; use of the conventional TBM leads to significant differences compared to the benchmark NEGF data, as illustrated by the results in Figs. 3.4 and 3.5.

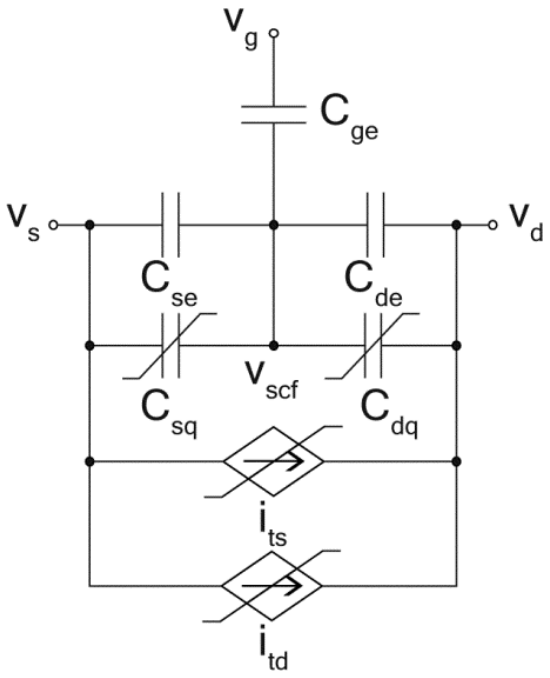
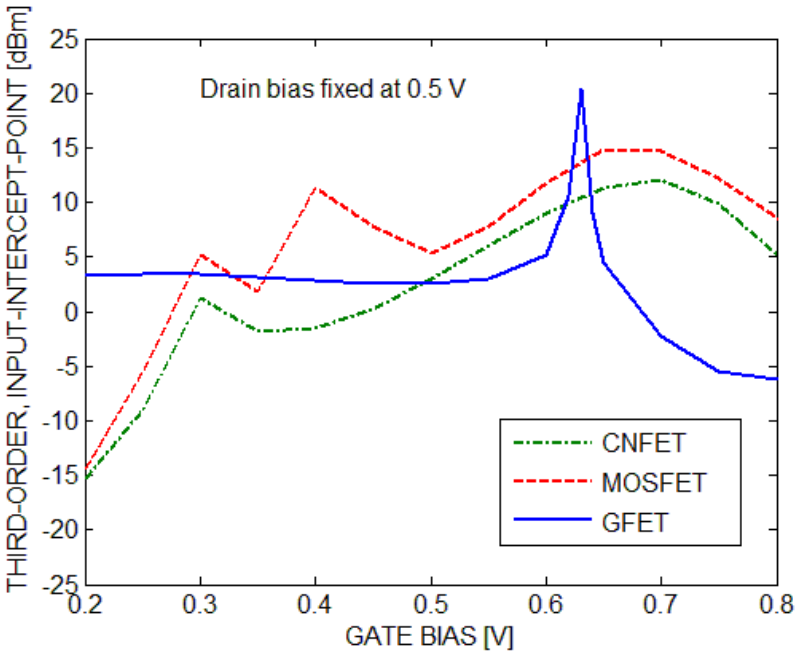
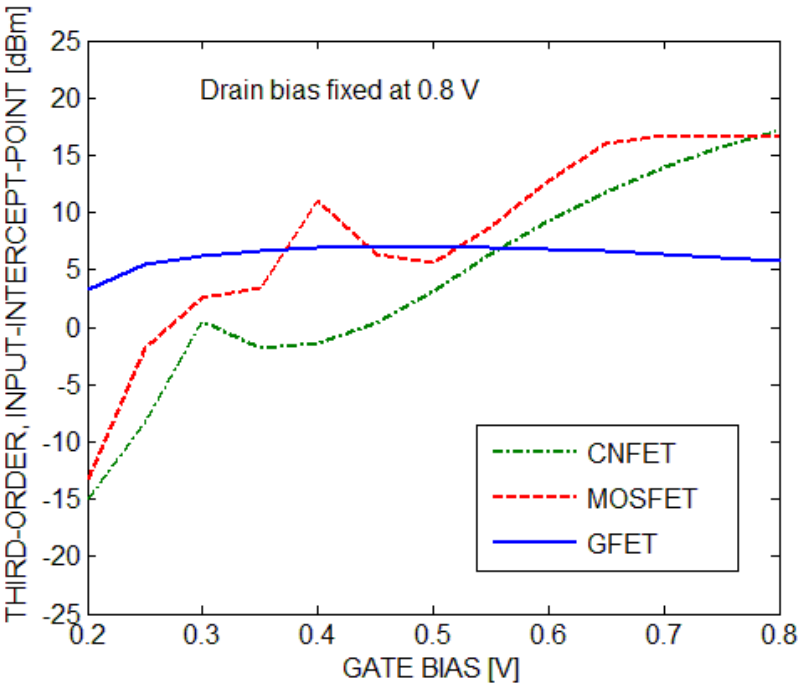


Fig. 3.6. Nonlinear small-signal equivalent circuit of a ballistic GFET.

Once the circuit is formed, the HB solver in Microwave Office [30] can be used to determine the GFET's third-order, input-intercept point (IIP3). Fig. 3.7(a) shows that for a drain bias of 0.5 V, the IIP3 of the GFET is comparable to its CNFET and MOSFET counterparts, where the latter have dimensions and oxide capacitances identical to the GFET. Increasing the drain bias to 0.8 V slightly improves the GFET linearity but the overall IIP3 remains comparable to that of the CNFET and the MOSFET, as shown in Fig. 3.7(b). However, unlike a CNFET and a MOSFET, where the drain biasing has a negligible effect on the characteristics of the IIP3 curve, our results show that the shape of the IIP3 curve of the GFETs is quite sensitive to the drain bias. Further such insights into the linearity and RF behavior of graphene-based devices can be obtained by exploiting our modified top-of-the-barrier model as described in Chapter 4.



(a)



(b)

Fig. 3.7. IIP3 versus gate bias, *with the drain bias held fixed at (a) 0.5 V and (b) 0.8 V*, of the graphene FET compared to its CNFET and MOSFET counterparts.

3.4. Conclusions

We developed a modified top-of-the-barrier model for GFETs that includes variations in the reservoir versus channel densities of states and band-to-band tunneling. The model shows excellent agreement with state-of-the-art, quantum-mechanical approaches based on NEGF and allows for the development of accurate, practical circuit models. The utility of the new model is demonstrated by exploring the RF linearity of graphene transistors. A detailed investigation of the nonlinear behavior of GFETs based on the developed model is described in Chapter 4.

Chapter 4

RF Linearity Performance Potential of Short-Channel Graphene Field-Effect Transistors[9]

4.1. Introduction

The unique electrical and physical properties of graphene have sparked much interest in determining its potential uses in electronics. Although the lack of a bandgap has been problematic for the use of graphene in digital applications, the high values of f_T and $f_{\max}$, combined with a high carrier mobility, continue to make graphene a promising candidate for analog high-frequency, or RF, electronics. A key figure-of-merit for RF applications is linearity, which measures the degree of distortion generated by the nonlinear mixing of the input signal with jammers. This work probes the performance potential of graphene in terms of RF linearity.

The strong interest in graphene has resulted in many theoretical and experimental studies on GFETs. These studies have largely focused on cut-off frequencies [76-78], mobility [6, 79, 80], the effect of the lack of a bandgap [10], and ways to introduce a bandgap to improve performance [81-83]. GFETs operating at promisingly high frequencies have already been demonstrated [84]. Furthermore, great progress has been made in the pursuit of graphene-based integrated circuits [85, 86].

On the topic of graphene linearity, however, there has been limited experimental work,

[9] A version of this chapter has been submitted for publication [17].

which can be summarized as follows. Wang *et al*. [36] investigated the linearity of a 2-μm long single transistor RF mixer at 10 MHz and reported an IIP3 of ~13.8 dBm; however, the reported conversion loss was between 30 to 40 dB. Habibpour *et al.* [37] reported a mixer based on a 500-nm long multichannel GFET operating at 30 GHz, with IIP3 values as high as 12.8 dBm and a conversion loss of 19 dB. Andersson *et al.* [38] reported the linearity of subharmonic mixers based on resistive GFETs having a channel length of 1 μm; they obtained an IIP3 of 4.9 dBm and a conversion loss of 20 – 22 dB. The shortest channel GFET investigated for RF linearity thus far is a 250-nm epitaxially grown graphene FET used as a mixer, reported by Moon *et al*. [39] with an IIP3 of ~22 dBm; however, the gain was small and in the range of < -15 dB; they also reported a similar but longer channel (2 μm) device with higher IIP3 (~27 dBm) and a small but somewhat better gain of -10 dB. Madan *et al*. studied the linearity of an RF mixer [40] and LNA [41] based on a 750-nm long graphene FET and reported third-order output-intercept point (OIP3) values in the range of 19 dBm at an operating frequency of 2 GHz; the gain of the LNA for a 50-Ω load termination was -5 dB. Jenkins *et al*. [42] also reported relatively good linearity for graphene FETs containing channels grown both by chemical vapor deposition and epitaxy and having lengths above 500 nm, with IIP3 values as high as 20 dBm but a power gain of < 15 dB for a 50-Ω load at 300 MHz.

Common trends in the results cited above are a long channel length (> 250 nm) for the devices and promising values of IIP3 that are accompanied by very low power gains. Further investigation is thus necessary to fully understand the RF linearity potential of GFETs, particularly the linearity that could be realized at short channel lengths.

Given the present difficulty of fabricating GFETs with channel lengths at or below those for current CMOS technology nodes, studying the linearity of GFETs with short channels, *i.e.*, $\lesssim 20$ nm, which is 10 to 100 times smaller than the reported experimental devices cited above, calls for a modeling approach. To date, there have been only a few modeling studies that explore GFET linearity, and these have also focused on longer channel devices ($\geq$ 440 nm). Chauhan *et al.* [31] have used a semi-classical model incorporating the effects of inelastic phonon scattering and have reported "excellent" linearity; however, the claim was based solely on the fact that the transconductance of their (1-μm long) GFET was observed to remain nearly constant over a wide range of

gate bias. Parrish *et al.* [32] have reported an analytical study in which they point out that the contact resistances can have a very detrimental effect on GFET transconductance linearity; working on a 2.4-μm long device, they showed that the IIP3 of a GFET can improve by as much as 17 dB if the contact resistances are made small enough to be neglected. Very recently, Rodriguez *et al.* [33] have used a static (low-frequency) analytical model to investigate the transconductance linearity of a 440-nm long GFET and have reported a peak IIP3 value of 13.8 dBm. None of these modeling studies accounted for all sources of nonlinearity relevant for RF performance, and in particular, transport *and* capacitive nonlinearities, both of which can be expected to play a role [15]. A detailed and more comprehensive study of the RF linearity mechanisms of short-channel GFETs is thus warranted.

In this work, we provide insight into the linearity mechanisms of an 18-nm GFET, chosen for demonstration purposes and representative of current CMOS technology nodes [9]. As in [27], we assume ballistic transport, a reasonable first approximation for graphene at small channel lengths ($\lesssim$ 20 nm) for the purposes of assessing performance potential, especially since the reported electron mean-free path in graphene is much larger ($\gtrsim$ 100 nm) [7]. We also consider a doped MOSFET-like device, as done in recent studies to assess the performance potential of carbon-based electronics [14, 27]; short-channel MOSFET-like devices can be expected to outperform the long-channel Schottky-barrier devices prevalent today [35] and are a suitable choice to gauge performance potential.

We start our analysis by using an already developed nonlinear small-signal circuit [15]. The intrinsic components of the circuit are first extracted based on a modified top-of-the-barrier model (MTBM) [16]. The MTBM is an extension of the conventional top-of-the-barrier model [14, 27], with additional features to account for physical effects arising from the lack of a bandgap in graphene**;** for further details, the reader is referred to [16]. The external parasitics are then calculated with the aid of COMSOL [34] and added to obtain a complete extrinsic nonlinear circuit, an approach which has already been shown [16] to capture the nonlinear voltage dependencies of key device parameters determined from a more detailed simulator [10]. The HB solver in Microwave Office (MWO) [30] is then used to simulate the developed nonlinear circuit.

Based on an examination of IIP3 values for intermodulation distortion under a two-tone input, our study reveals that a GFET's linearity has a signature behavior versus gate bias that includes a constant-linearity region at low gate bias, sweet spots of high linearity before and after the gate bias for peak cutoff frequency, and poor linearity at the gate bias corresponding to the peak cutoff frequency. We otherwise find that a GFET offers overall linearity that is comparable to a MOSFET and a CNFET, with the exception that the amount of intermodulation distortion in a GFET is dominated by the drain-injected carriers, a unique outcome of graphene's lack of a bandgap. We also examine the effects of drain bias, load resistance, and external parasitics.

Section 4.2 outlines the device structure and simulation methodologies. The results are presented and discussed in Section 4.3, and a qualitative comparison of these results with experiment is provided in Section 4.4. The conclusions of this study are summarized in Section 4.5.

4.2. Approach

4.2.1. Device Structure

Fig. 4.1 shows the schematic of the GFET under investigation, with key device dimensions marked. The dotted region indicates the intrinsic portion of the device. The gate oxide is a 2-nm layer of Al_2O_3 (with a relative permittivity $\epsilon_r = 9.8$). Al_2O_3 has been demonstrated as a promising high-k dielectric suitable for graphene in recent experiments [87, 88]. The channel is intrinsic graphene, while the source and drain regions are n-doped, with an effective doping concentration of $N_D = 1.9 \times 10^{17}$ m^{-2}. The source and drain geometries are symmetric with respect to the channel/gate regions.

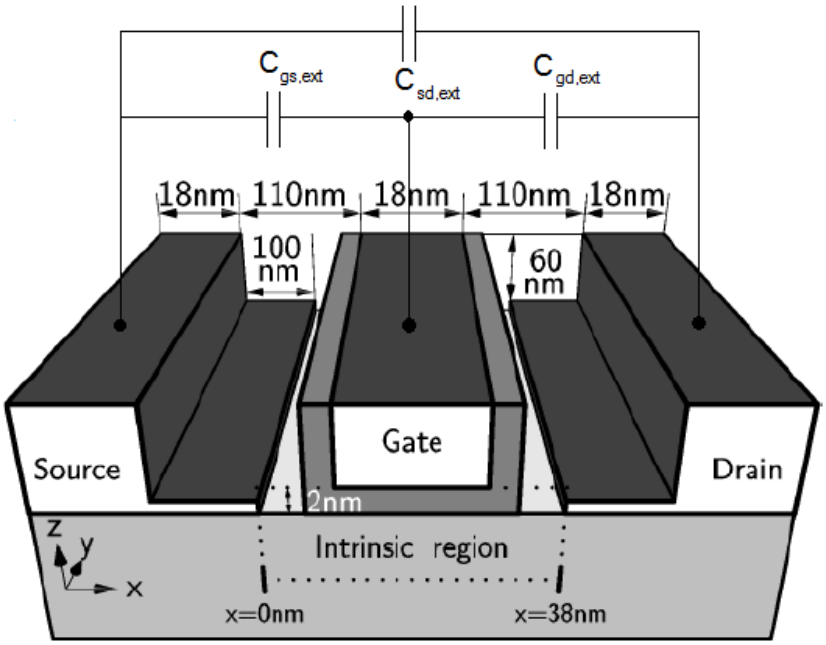


Fig. 4.1. Schematic of the GFET used in this work. The dotted intrinsic region is modeled using a modified top-of-the-barrier method [16]. The external parasitic capacitances used to model the extrinsic device are shown at the top of the schematic.

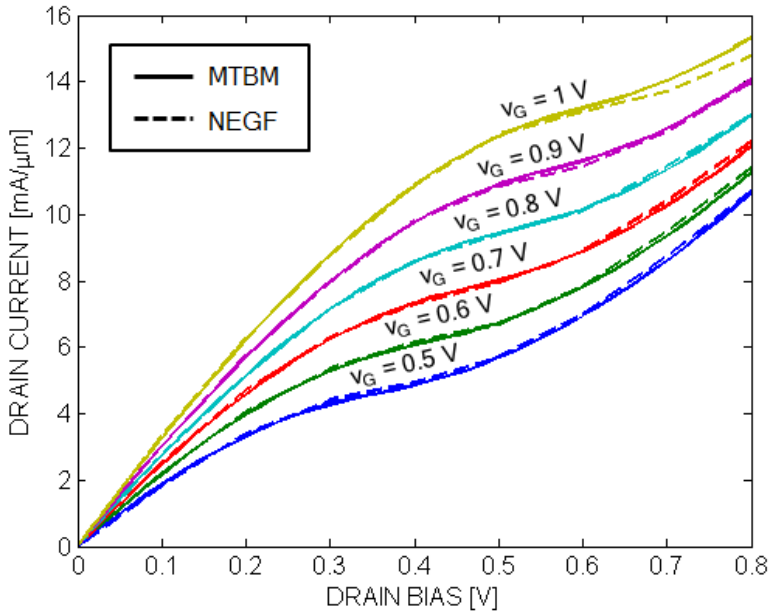


Fig. 4.2. Current-voltage characteristics of the GFET under investigation from MTBM and NEGF. Reproduced from [16].

Fig. 4.2 plots the current-voltage characteristics of the GFET calculated using the MTBM [16] and a fully quantum-mechanical solver based on NEGF [10]. Other important parameters such as transconductance g_m, output conductance g_{o} (or g_{ds}), and unity-current-gain (cutoff) frequency f_T of the device were reported in [16]; plots of these quantities versus gate bias are also available in figures (4 and 11) discussed further below.

4.2.2. Intrinsic Equivalent Circuit

Our focus in this work is the small-signal nonlinear operation of GFETs. We hence use Taylor-series expansions for all the components in the small-signal equivalent circuit. The coefficients of the series are specified by derivatives [evaluated at corresponding bias (dc operating) point] of the charge-voltage and current-voltage relationships from the MTBM [16]. The dotted portion of Fig. 4.3 represents the intrinsic nonlinear small-signal equivalent circuit, where the elements are as follows: C_{ge}, C_{se}, and C_{de} are the linear electrostatic capacitances of the GFET; C_{sq} and C_{dq} are the nonlinear source and drain quantum capacitances; and i_{ts} and i_{td} are the nonlinear current sources modeling the quasi-static transport currents of the device. Each of the nonlinear components are represented by a Taylor-series expansion up to third order:

$$q_{\mathrm{sq}} = C_{\mathrm{sq1}}(v_s - v_{\mathrm{scf}}) + C_{\mathrm{sq2}}(v_s - v_{\mathrm{scf}})^2 + C_{\mathrm{sq3}}(v_s - v_{\mathrm{scf}})^3 \tag{4.1}$$

$$q_{\mathrm{dq}} = C_{\mathrm{dq1}}(v_d - v_{\mathrm{scf}}) + C_{\mathrm{dq2}}(v_d - v_{\mathrm{scf}})^2 + C_{\mathrm{dq3}}(v_d - v_{\mathrm{scf}})^3 \tag{4.2}$$

$$i_{\mathrm{ts}} = g_{\mathrm{sq1}}(v_s - v_{\mathrm{scf}}) + g_{\mathrm{sq2}}(v_s - v_{\mathrm{scf}})^2 + g_{\mathrm{sq3}}(v_s - v_{\mathrm{scf}})^3 \tag{4.3}$$

$$i_{\mathrm{td}} = g_{\mathrm{dq1}}(v_d - v_{\mathrm{scf}}) + g_{\mathrm{dq2}}(v_d - v_{\mathrm{scf}})^2 + g_{\mathrm{dq3}}(v_d - v_{\mathrm{scf}})^3 \tag{4.4}$$

where q_{sq} and q_{dq} are the small-signal (ac) parts of the charges held by the quantum capacitances C_{sq} and C_{dq}, respectively; v_s and v_d are the small-signal parts of the source and drain voltages, respectively; and v_{scf} is the small-signal part of the (self-consistent)

channel potential. The steps described in [15] were followed to determine the values of the linear and nonlinear components from the MTBM [16].

4.2.3. Extrinsic Equivalent Circuit

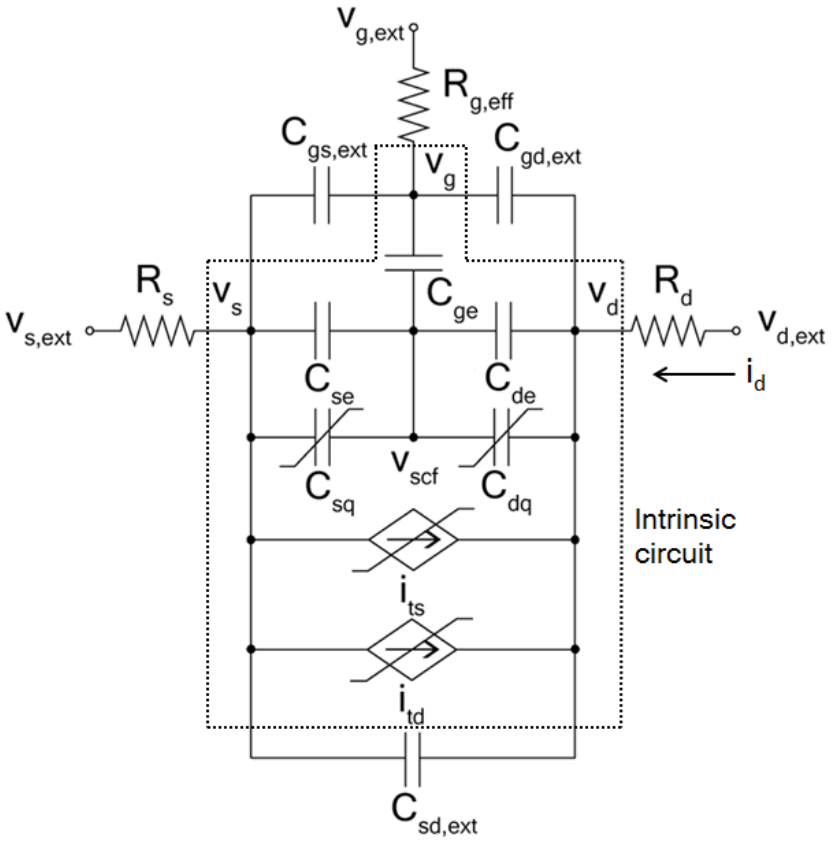


Fig. 4.3. Complete nonlinear small-signal equivalent circuit of a ballistic GFET.

The performance of a practical GFET is also impacted by the parasitic elements in the device due to the metallic contacts at the gate, source, and drain. In order to fully assess the linearity of these devices, the effects of these parasitics must be incorporated. We therefore add the extrinsic capacitances $C_{\mathrm{gs,ext}}$, $C_{\mathrm{gd,ext}}$, and $C_{\mathrm{sd,ext}}$, labeled in Fig. 4.1 along with the contact resistances of the gate, drain, and source, $R_{\mathrm{g,eff}}$, R_s, and R_d, respectively. All the parasitic components were calculated following the method described in [15] with the aid of COMSOL [34], and by using the contact dimensions specified below in Section 4.3.4(a). The resulting extrinsic nonlinear small-signal

equivalent circuit is the overall circuit in Fig. 4.3, where v_g, v_s, and v_d are the internal node voltages of the GFET and $v_{\mathrm{g,ext}}$, $v_{\mathrm{s,ext}}$, and $v_{\mathrm{d,ext}}$ are the external terminal voltages of the overall device. The component values (both intrinsic and extrinsic) are listed in Table 4.2 in Section 4.3.4(a) for the device under investigation.

4.3. Results and Discussion

We used the HB solver in MWO [30] to simulate the nonlinear small-signal equivalent circuit, and we extracted the IIP3 corresponding to the mixing frequency $2f_1 - f_2$, under excitation from two input tones at the fundamental frequencies f_1 and f_2, as the small-signal linearity figure of merit of the device. The transistor was deployed in a simple common-source configuration. The load and source impedances were set at $50\ \Omega$, the usual characteristic impedance for RF applications. A two-tone source with an impedance of $50\ \Omega$ and an operating frequency of 24 GHz – which is a frequency of interest in RF electronics according to the 2012 ITRS [9] – and a difference of 100 kHz between the two tones was used ($f_1 = 24$ GHz, $f_2 = 24.1$ GHz). The source was grounded, the drain bias was fixed at $v_D = 0.5$ V, and the gate bias was varied over a wide range, from 0 to 1 V (except for the results in Figs. 4.6 and 4.10, where the upper limits are 1.2 V and 1.4 V, respectively, and Fig. 4.19 where the range is from -0.2 V to 1.2 V, to aid the discussion).

4.3.1. Key Features of GFET Linearity

To reveal the key features of GFET linearity, we first investigated the intrinsic RF linearity of a GFET, *i.e.*, the linearity determined by the dotted portion of Fig. 4.3 and excluding external parasitics. The resulting IIP3 was plotted against variations in gate bias and is shown in Fig. 4.4. The IIP3 curve has a very distinct shape (signature), with a constant linearity region (region 1), two sharp peaks at points 2 and 4, and a large dip around point 3. The presence of the peaks at points 2 and 4 mean that bias sweet spots may exist where a GFET will behave very linearly. Fig. 4.4 also shows the unity-current-

gain frequency f_T versus gate bias. Note that the peak f_T coincides with point 3, which means the GFET is most nonlinear at peak f_T.

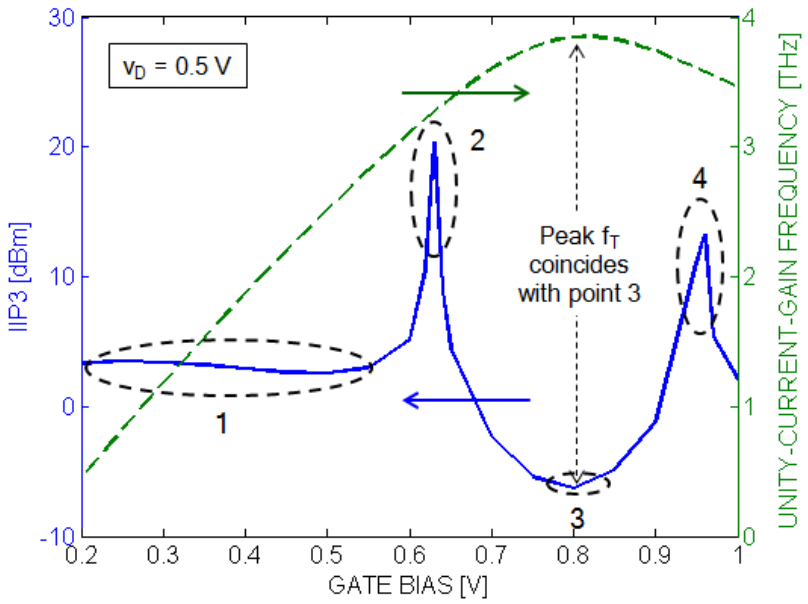
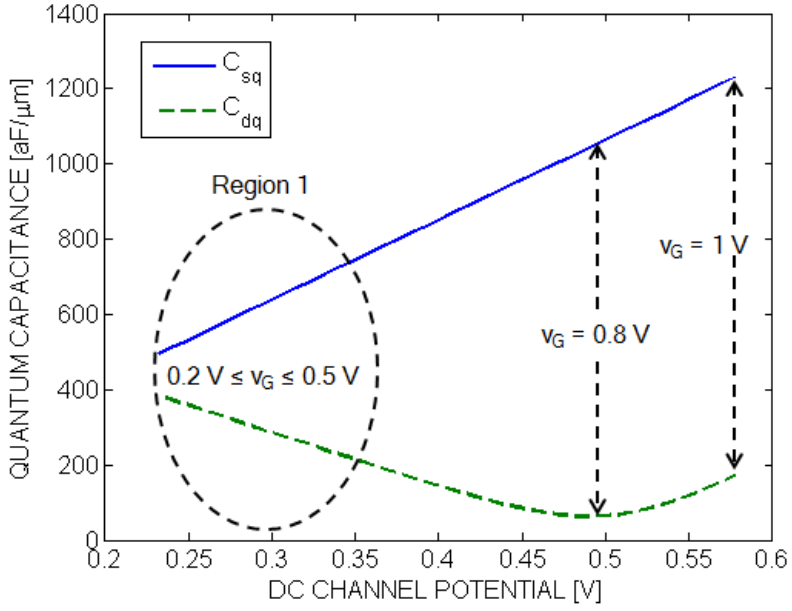


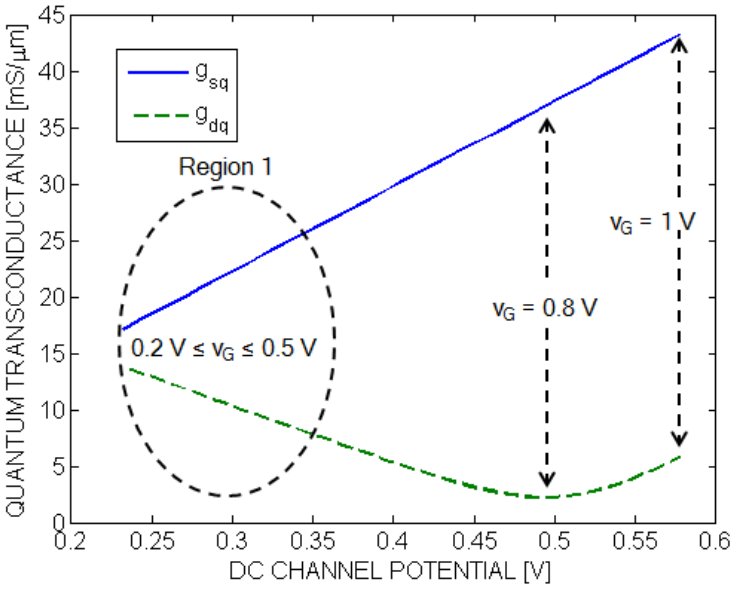
Fig. 4.4. Intrinsic IIP3 and unity-current-gain frequency f_T versus gate bias for the GFET under investigation.

a) Constant IIP3 Region (Region 1)

From the small-signal equivalent circuit in Fig. 4.3, it is clear that the distortion in a GFET arises from the nonlinear quantum capacitances and current sources, labeled C_{sq}, C_{dq}, i_{ts}, and i_{td}. More precisely, intermodulation distortion at the third-order mixing frequency $2f_1 - f_2$, which is of principal interest in this work, arises due to the second- and third-order coefficients of the corresponding Taylor series expansions (4.1)-(4.4) for these elements; the second-order coefficients contribute by creating second-order distortion and then re-mixing it with the fundamental frequencies, and the third-order terms contribute by directly mixing the fundamental frequencies. We hence focus our attention on the behavior of both the second- and third-order coefficients.



(a)



(b)

Fig. 4.5. (a) Quantum capacitance and (b) transconductance versus the dc part of the channel potential for the GFET under investigation. The gate bias voltages for a few points are indicated for reference.

Fig. 4.5 plots the quantum capacitances (C_{sq}, C_{dq}) and quantum transconductances (g_{sq}, g_{dq}) with respect to the bias (dc) part of the channel potential v_{SCF}, where the capacitances are defined as the derivatives of the source- and drain-injected charge with respect to the channel potential, respectively, and the transconductances are similarly defined but involving derivatives of currents [14]. The values of gate bias voltage v_G that apply are also indicated for a few points on the plots. Since the curves were obtained with constant source and drain voltages, then by definition, the values of the capacitances (C_{sq}, C_{dq}) and transconductances (g_{sq}, g_{dq}) on the plots are the *first*-order coefficients C_{sq1}, C_{dq1}, g_{sq1}, and g_{dq1} appearing in (4.1)-(4.4). The second- and third-order coefficients in (4.1)-(4.4) are therefore determined by the first and second derivatives of the curves in Fig. 4.5. It can be seen that in region 1 ($0.2\ \mathrm{V} \leq v_G \leq 0.5\ \mathrm{V}$), the curves vary linearly with voltage, which means that the third-order coefficients (determined by the second derivatives) are almost zero, while the second-order coefficients (first derivatives) are constant, thereby yielding a steady amount of distortion in the device over region 1. The linear behavior of the capacitances and transconductances in region 1, and hence the constant IIP3 in region 1, arise from the linear density of states (DOS) of graphene; the connection between the DOS and the expected behavior is explained for C_{dq} and g_{dq} when discussing Fig. 4.8, and similar reasoning applies for C_{sq} and g_{sq}.

b) Sharp Peaks at Points 2 and 4

The distortion in a GFET can arise from multiple sources, and the distortion generated from these sources can act upon each other constructively or destructively. In the discussion to follow, we show that the peaks at points 2 and 4 arise due to the destructive combination of distortion from two different sources. Using appropriate biasing, it may therefore be possible to make the GFET behave very linearly.

To identify how the GFET's linearity is affected by the contributions from the second- and third-order coefficients in (4.1)-(4.4), we turn them on and off selectively in the intrinsic equivalent circuit of Fig. 4.3. Fig. 4.6 shows the IIP3 of the GFET due to the two types of coefficients.

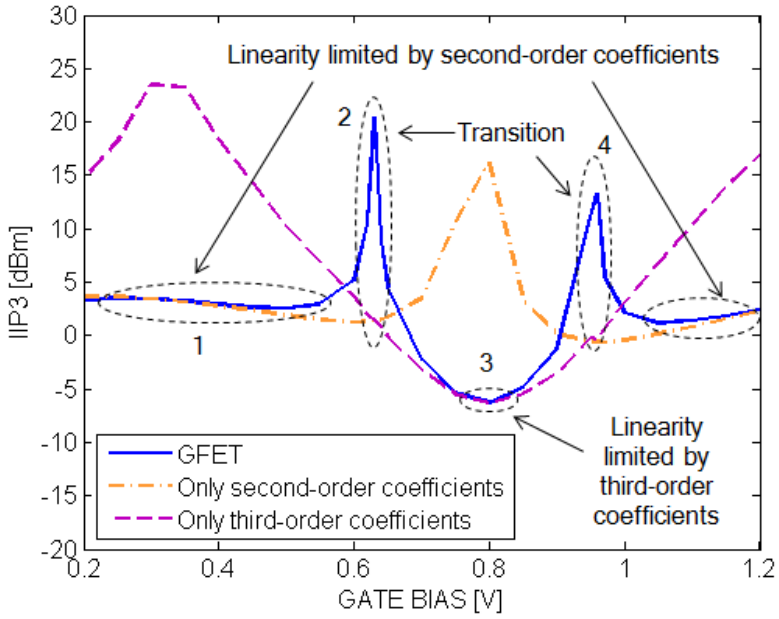


Fig. 4.6. Effects of the second- and third-order coefficients on the overall linearity of the GFET.

It is evident that over particular bias points, the GFET linearity is determined by one of the two types of coefficients. At low and high gate biases ($v_G < 0.6$ V and $v_G > 1.0$ V), the device linearity is limited by the distortion generated by the second-order coefficients. However, for the moderate bias range 0.7 V $\leq v_G \leq$ 0.9 V, the device linearity is limited by distortion generated by the third-order coefficients. The peaks at points 2 and 4 appear when the device linearity mechanism switches from one type to the other. These results strongly suggest that at the transition regions, distortion contributions from the two mechanisms are combining in such a way that they cancel each other, making the device extremely linear. To illustrate the cancellation, MWO was used to generate the distortion components of the small-signal output current i_{d}, at the mixing frequency $2f_1 - f_2$, in the transition regions; Fig. 4.7 shows the results at a gate bias of 0.63 V (point 2). It is seen that the distortion due to the two mechanisms (second- and third-order coefficients) are indeed 180° out of phase. Similar behavior is observed at the gate bias of 0.96 V (point 4).

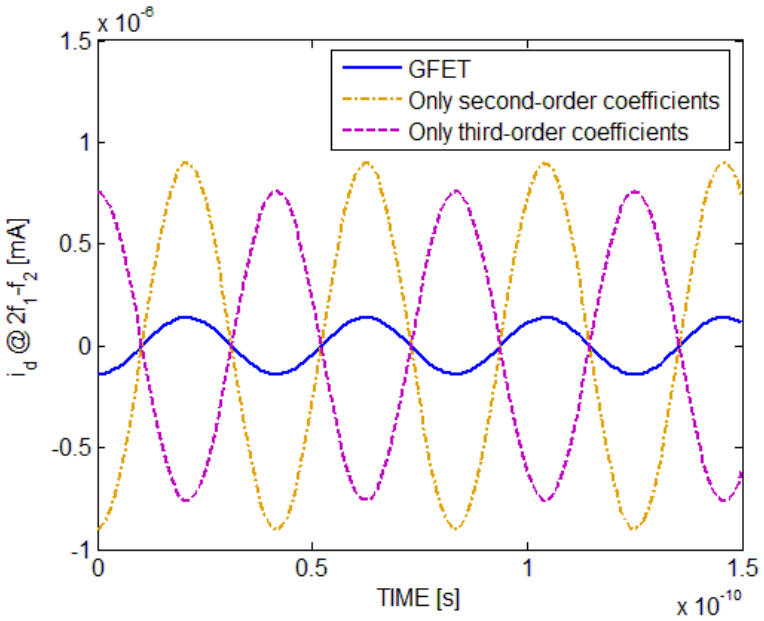


Fig. 4.7. Distortion components at $2f_1 - f_2$ in i_{d} at a gate bias of 0.63 V (point 2 in Fig. 4.6). A destructive combination of the distortion from the two types of sources results in a diminished overall distortion.

c) Dip at Point 3

The dip in GFET IIP3 at point 3 occurs where device linearity is limited by the third-order coefficients, as shown by the results in Fig. 4.6. An inspection of Fig. 4.5 shows that the source components C_{sq} and g_{sq} [solid lines in parts (a) and (b) of Fig. 4.5] behave linearly over all gate biases of interest, meaning their third-order coefficients (determined by the second derivatives) are zero, but the drain components C_{dq} and g_{dq} [dashed lines in parts (a) and (b) of Fig. 4.5] both show minima at point 3 ($v_G = 0.8$ V), which leads to large third-order coefficients (determined by the second derivatives). The nonlinear elements C_{dq} and g_{dq} associated with the drain can hence be expected to contribute substantial distortion around point 3, which limits the device linearity, as illustrated in Fig. 4.6. The origin of the minima in C_{dq} and g_{dq} can be explained by observing what happens to the drain transport in this bias region.

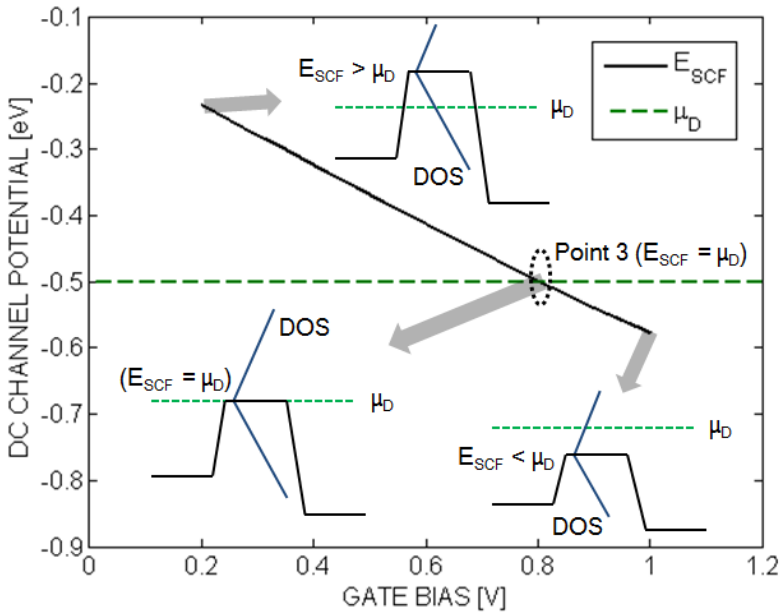


Fig. 4.8. The dc part of the channel potential $E_{\mathrm{SCF}} = -qv_{\mathrm{SCF}}$ versus gate bias. The drain Fermi level μ_D (fixed by the constant drain voltage of 0.5 V) is also shown. The source Fermi level μ_S (not shown) is taken to be the reference ($\mu_S = 0$ eV).

Fig. 4.8 shows the dc channel potential $E_{\mathrm{SCF}} = -qv_{\mathrm{SCF}}$, equivalent to the position of the Dirac point in the channel of a GFET, as a function of gate bias. As illustrated, the channel potential (Dirac point) decreases with an increasing gate bias and crosses the drain Fermi level $\mu_D = -0.5$ eV at point 3, *i.e.*, for $v_G = 0.8$ V. The insets in Fig. 4.8 are provided as visualization aids and show the position of the Dirac point and channel DOS with respect to the drain Fermi level at a few gate biases. It can be seen that at lower gate biases ($v_G < 0.8$ V), μ_D is positioned below the channel potential and a large number of states are available in the channel at the drain Fermi level. As the gate bias increases, the available DOS at μ_D starts to decrease and becomes zero at point 3, where the channel Dirac point aligns with the drain Fermi level ($E_{\mathrm{SCF}} = \mu_D$). Beyond point 3 ($v_G > 0.8$ V), μ_D is positioned above the channel potential, and the number of states available at the drain Fermi level increases with gate bias. Since the drain quantum capacitance C_{dq} depends directly on the available DOS at the drain Fermi level [66], it follows the same trend, *i.e.*, C_{dq} decreases linearly with gate bias before reaching point 3, becomes a minimum at point 3, and increases linearly after point 3. The (energy-

independent) constant velocity of electrons (and holes) in graphene means that i_{td} in Fig. 4.3 and its first derivative g_{dq} behave in the same way as q_{dq} and its first derivative C_{dq}, respectively, which can be discerned by their governing equations [14]. Thus both C_{dq} and g_{dq} show minima at $v_G = 0.8$ V (Fig. 4.5), *i.e.*, at point 3 (Fig. 4.6).

4.3.2. Drain Dominance in GFET Linearity

To further investigate the role of the drain in determining the linearity of a GFET, we selectively turned on and off the distortions from the source and drain components, by setting the appropriate higher-order coefficients in (4.1)-(4.4) to zero. The results are shown in Fig. 4.9. For all gate biases, the linearity of the device is found to be dominated by distortion coming from the drain. This result is significantly different from a conventional field-effect transistor in which the channel material has a finite bandgap (MOSFET or CNFET), where the distortion primarily comes from the source components [15]. The reason behind this unique drain dependency of the GFET linearity is two-fold:

i) The zero bandgap of graphene means that the drain always contributes to the transport. Consequently, the drain quantum capacitance and quantum transconductance of the GFET [dashed curves in Figs. 4.5(a) and 4.5(b)] are always large enough to impact the overall device behavior. Fig. 4.5 shows that the drain components C_{dq} and g_{dq} of quantum capacitance and quantum transconductance are relatively large, *i.e.*, of the same order as the source components, and that they are nonlinear in a GFET, unlike a conventional MOSFET (where C_{dq} and g_{dq} are essentially zero [15]). Comparing the capacitance-voltage and transconductance-voltage relationships of the source and drain components in Fig. 4.5, it is evident that the resulting second-order coefficients in (4.1)-(4.4), determined by the first derivatives of the curves, would be comparable. On the other hand, the minima in C_{dq} and g_{dq} make the third-order coefficients (determined by the second derivatives) of the drain components much larger than the almost zero third-order coefficients of the

source components. The drain components can thus be expected to produce more distortion than the source components in a GFET.

ii) The common-source configuration of the device makes the small-signal gain negative, which means that the small-signal drain voltage v_d is 180° out of phase with the small-signal gate voltage v_g, and hence with the small-signal channel potential v_{scf} (which will tend to follow v_g). This phase difference makes the control voltage $(v_d - v_{\text{scf}})$ for the drain components of Fig. 4.3, governed by (4.2) and (4.4), bigger than the corresponding control voltage v_{scf} for the source components, governed by (4.1) and (4.3). The larger control voltage enhances the distortion coming from the drain components.

The following discussion highlights some of the outcomes of this unique drain dominance in GFET linearity.

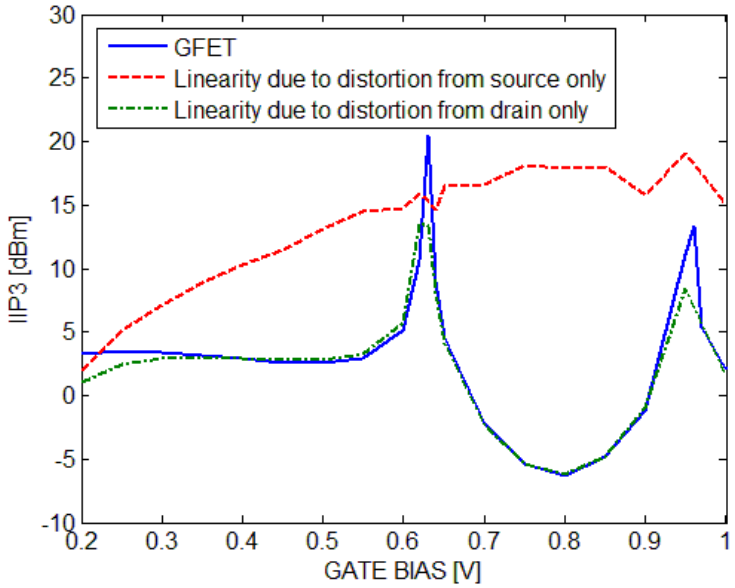


Fig. 4.9. Effect of distortion from the source and drain on IIP3. Linearity due only to the source was found by setting the higher-order coefficients in (4.2) and (4.4) to zero; similarly, linearity due only to the drain was found by setting the higher-order coefficients in (4.1) and (4.3) to zero.

a) Effect of Drain Bias on Linearity

One obvious outcome of the drain dominance on GFET linearity is an expected drain-bias dependency of the overall linearity. Fig. 4.10 shows the IIP3 of the GFET versus gate bias, at a few different values of drain bias.

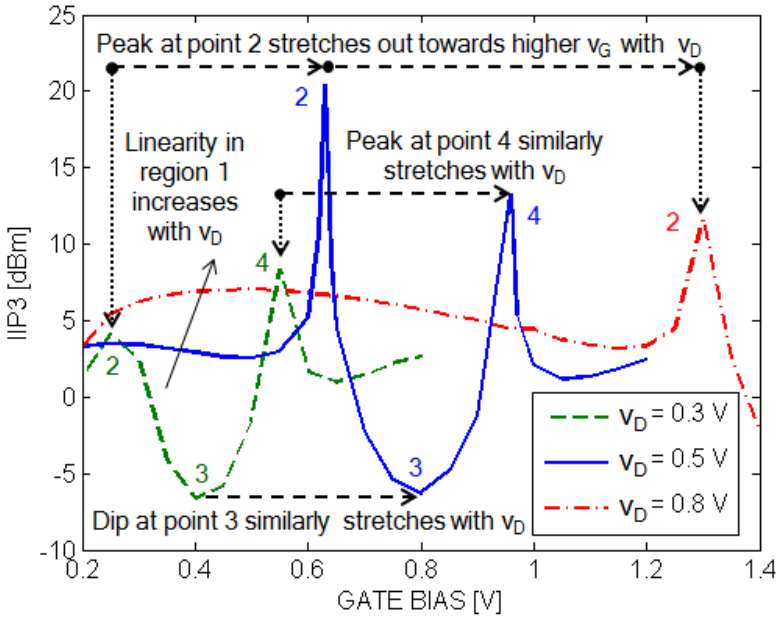
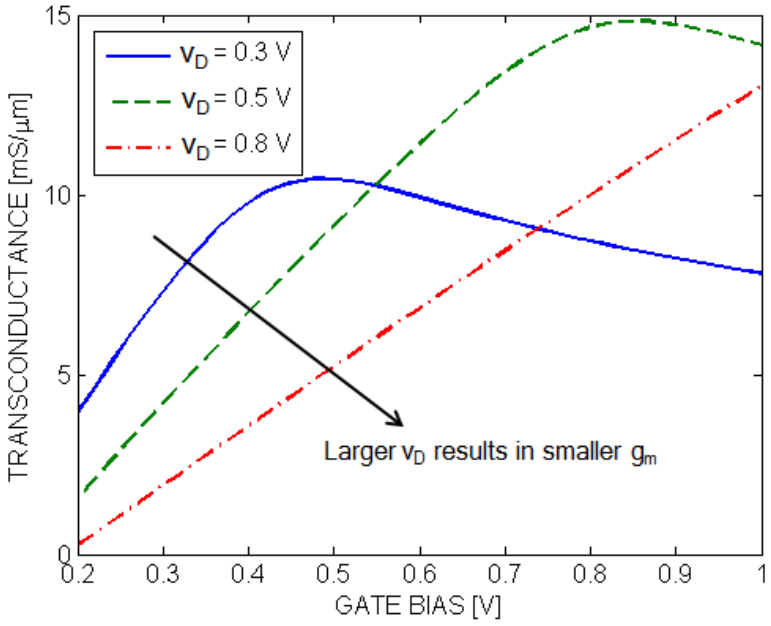


Fig. 4.10. IIP3 versus gate bias, at a few different values of drain bias.

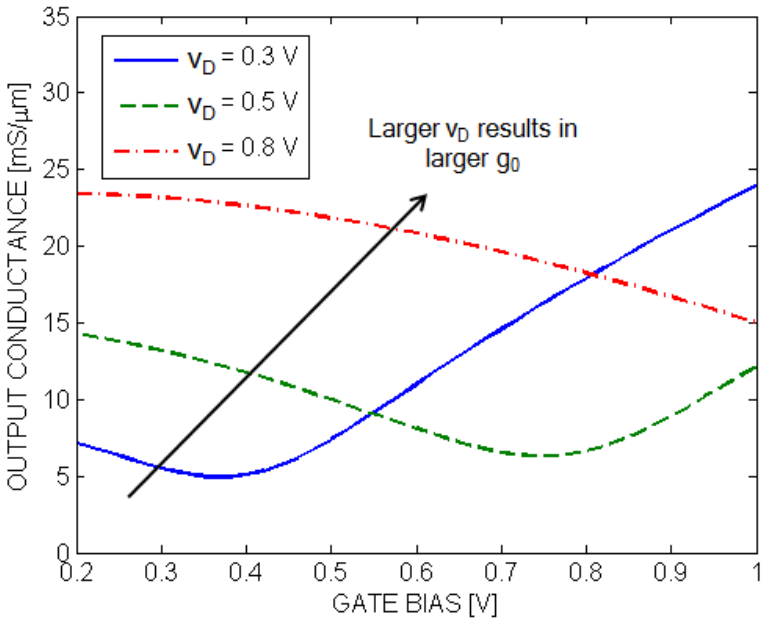
As can be seen from the figure, in region 1 (constant linearity), a larger drain bias makes the device more linear. This outcome can be explained with the help of Fig. 4.11, which shows that in this region, a larger drain bias reduces the transconductance g_m and increases the output conductance g_o. Since the available small-signal voltage gain of the GFET depends on the ratio g_m/g_o, the larger drain bias results in a smaller small-signal voltage gain and hence a smaller v_d. A reduced v_d means that the control voltage $(v_d - v_{scf})$ for the drain components is also reduced, which can be expected to reduce the distortion from the drain components in Fig. 4.3 [according to (4.2) and (4.4)] and make the GFET more linear.

A larger drain bias also *stretches* the IIP3 curve, pushing the peaks at points 2 and 4, along with the dip at point 3, towards higher gate biases. The straightforward reason for

this outcome is that a larger gate bias is required to push the dc channel potential (Dirac point on a band diagram) down to the lower drain Fermi level at higher drain bias.



(a)



(b)

Fig. 4.11. (a) Transconductance and (b) output conductance versus gate bias for varying drain bias.

b) Effect of Load Resistance on Linearity

Another outcome of the drain dominance on GFET linearity is the effect of the load resistance R_L. A larger R_L results in a larger swing in the drain voltage v_d, which enhances the amount of distortion from the drain components through a larger control voltage $(v_d - v_{\text{scf}})$ in (4.2) and (4.4). On the other hand, a smaller R_L results in a smaller swing in v_d and the distortion becomes smaller. Fig. 4.12 shows the effect of R_L on IIP3, while the source resistance is held at 50 Ω. As anticipated, reducing the load from 50 Ω to 12.5 Ω dramatically increases the GFET IIP3 by almost 10 dB. Similarly, increasing the value of R_L degrades linearity.

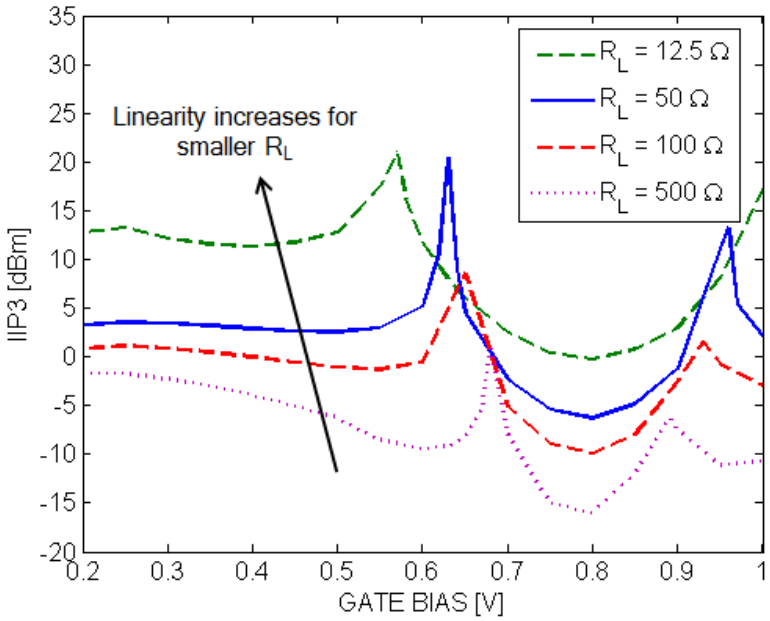


Fig. 4.12. Effect of load resistance R_L on GFET IIP3.

The unique zero bandgap of graphene (the reason behind the drain dominance) thus makes it possible to improve the linearity, by reducing the load resistance. However, before reducing R_L to improve the linearity, one must consider its implications on the voltage and power gains of the device, two desirable properties of any FET operating at RF frequencies.

(1) Voltage Gain

The large output conductance of a GFET limits the voltage gain achievable from these devices. For example, Fig. 4.11 shows that for a drain bias of 0.5 V, the maximum (open-circuit) voltage gain available from the GFET is $g_m/g_o \approx 1.2$ V/V at a gate bias of 0.5 V. The voltage gain becomes even smaller when the device is loaded with a finite R_L. Table 4.1 shows that the small R_L of 12.5 Ω that makes the GFET very linear in Fig. 4.12 also reduces the voltage gain to a mere 0.1 V/V. An attempt to improve linearity by reducing R_L thus reduces the voltage gain considerably.

(2) Power Gain

Even though the voltage gain of graphene is poor, a sufficiently wide device can still provide enough power gain (through increased current drive). For example, Table 4.1 shows that the 1-µm wide device has a power gain of -3.77 dB with a resistively matched load of 100 Ω, but a 10-µm wide device has a power gain of 6.23 dB with a resistively matched load of 10 Ω, where the degree of matching is indicated by the product $g_o R_L$. The power gain of the device can thus be increased by making the device wider and setting $g_o R_L = 1$. For simplicity, here we are discussing the power gain simply as $P_G = P_L/P_{av,s}$, where P_L is the power delivered to the load and $P_{av,s}$ is the power available from the source, under the condition of purely resistive terminations for which we have been examining the IIP3; substantially more gain is available, as indicated, for example, by the maximum available gain (MAG), which is ~30 dB for the 1-µm device [10].

Our conclusions on the behavior of IIP3 are unaffected by device width (Table 4.1), as long as we compare IIP3 values for the same $g_o R_L$. Hence, we can now consider a wider device, providing more power gain, and consider again the tradeoff between load resistance and linearity. For example, for the 10-µm wide device, reducing the load from 10 Ω to 1.25 Ω will improve the IIP3 from -1.08 dBm to 12.78 dBm, but will decrease the power gain from 6.23 dB to 2.35 dB.

While admittedly examined under highly simplified conditions (resistive terminations and for an intrinsic device), the key point from this discussion is that the results in Fig. 4.12 and Table 4.1 demonstrate that a reduction in R_L does have the potential to improve

linearity in a GFET, unlike conventional FETs, subject to the caveats of reduced voltage and power gain. We will re-examine this issue when external parasitics are introduced (Section 4.3.4 below).

TABLE 4.1. EFFECT OF CHANNEL WIDTH AND LOAD IMPEDANCE ON LINEARITY AND GAIN @ 24 GHZ AND 0.5 V OF GATE AND DRAIN BIAS

Channel Width (μm)	R_L (Ω)	$g_o R_L$	IIP3 (dBm)	Voltage Gain (V/V)	Power Gain (dB)
1	12.5	0.125	12.78	0.1	-7.65
	50	0.5	2.59	0.3	-4.21
	100	1	-1.08	0.46	-3.77
10	1.25	0.125	12.78	0.1	2.35
	5	0.5	2.6	0.31	5.78
	10	1	-1.08	0.46	6.23

4.3.3. Linearity of a GFET versus a MOSFET and a CNFET

a) Third-Order Intermodulation Distortion

In order to determine if a GFET holds any promise in RF electronics in terms of linearity, we need to benchmark its performance against its competitors. As a basis for comparison, we simulated the linearity of a silicon MOSFET and an array-based CNFET with identical channel length ($L = 18$ nm), channel width ($W = 1$ μm), and gate capacitance; these are the devices illustrated in [Fig. 1, 15]. The CNFET had 100 tubes (tube pitch = 10 nm) in the channel to obtain a drive current comparable to the other devices. All three devices (including the GFET) were tested with 50-Ω two-tone sources and 50-Ω load terminations and the IIP3 values were recorded against gate bias. For the comparison, we retain the focus on the linearity of the intrinsic transistor so that the emphasis in our comparison is on differences arising from the channel material. Fig. 4.13 shows that the GFET offers linearity that is, overall, comparable to its MOSFET and CNFET counterparts under this scenario. However, two differences can be flagged. First,

as already discussed, the drain dependence of the GFET offers us with an opportunity to enhance its linearity by increasing the drain bias v_D or by lowering R_L, which is not possible in the other devices. Second, the GFET's linearity offers a sweet spot prior to and after peak f_T; these are the points 2 and 4 discussed earlier in conjunction with Fig. 4.4. In fact, the GFET offers its worst IIP3 at peak f_T, unlike the MOSFET and the CNFET, both of which offer their best IIP3 at peak f_T.

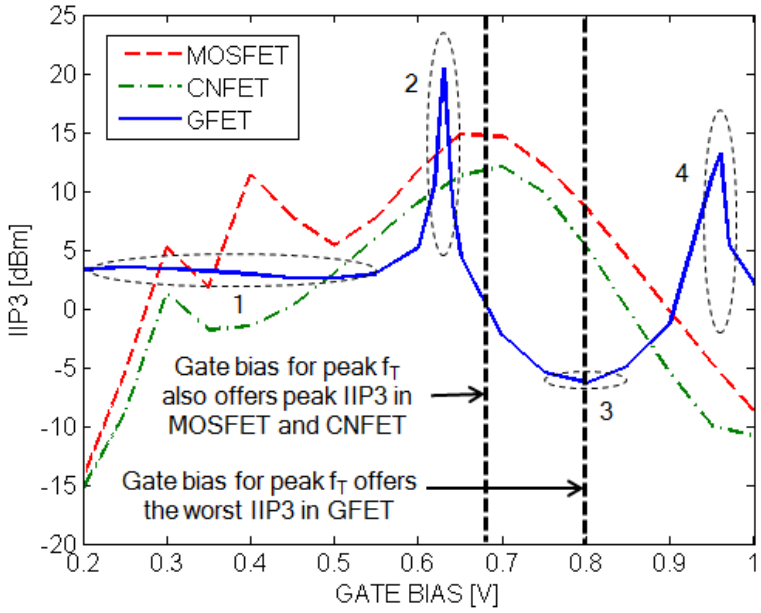


Fig. 4.13. Intrinsic linearity performance potential comparison of a GFET with its MOSFET and CNFET counterparts. The region 1 and points 2-4 from Fig. 4.4 for the GFET curve are marked. We have also indicated that the peak IIP3 for a CNFET and MOSFET occur at the same gate bias as peak f_T whereas for a GFET the minimum IIP3 occurs at the gate bias for peak f_T. The GFET f_T curve is available in Fig. 4.4; the MOSFET and CNFET f_T curves are not shown.

b) Second-Order Distortion

While we have focused on third-order distortion, second-order distortion can also be important in certain RF applications [89]. For example, two out-of-band jammers can mix via a second-order intermodulation product, creating undesired components at the

sum and difference frequencies, each of which could land on the fundamental frequency. We will focus on the sum frequency for the sake of this discussion.

For the second-order distortion at the mixing frequency $f_1 + f_2$, the GFET suffers from poor linearity when compared to its MOSFET and CNFET counterparts (Fig. 4.14). This outcome is primarily because of the linear DOS and zero bandgap of graphene, which cause all four quantum capacitances and transconductances of the GFET (source and drain components) to contribute to the distortion.

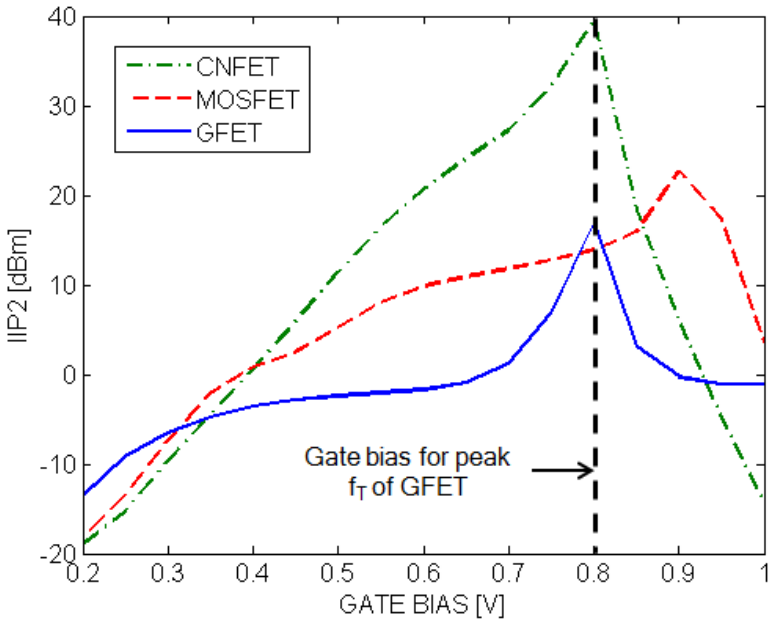
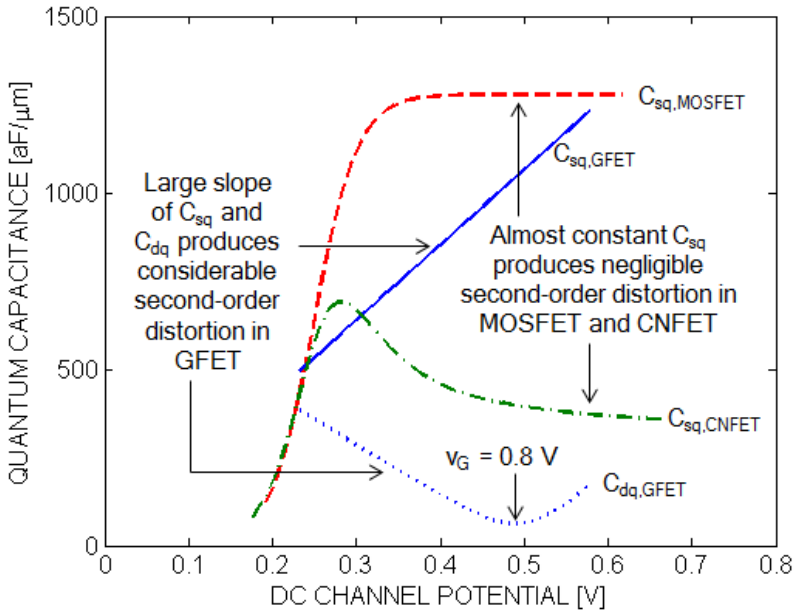
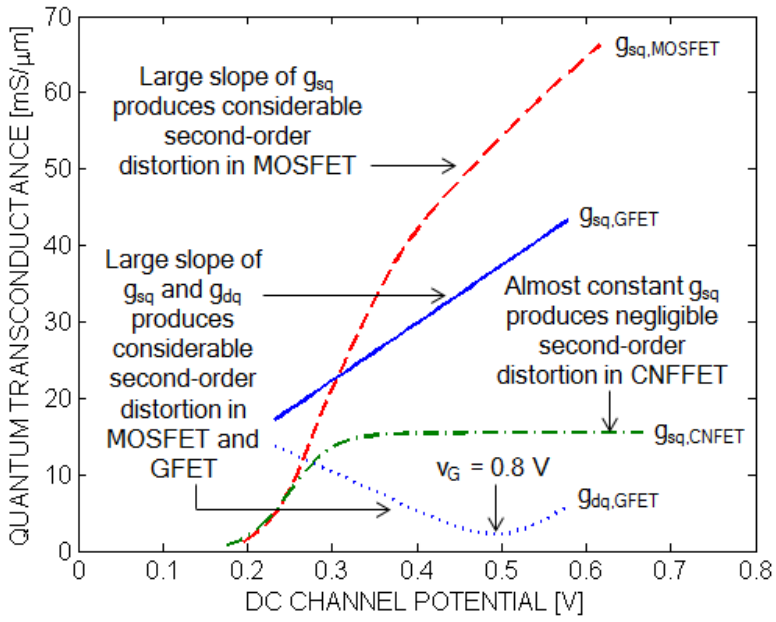


Fig. 4.14. Intrinsic IIP2 versus gate bias of a GFET compared with its MOSFET and CNFET counterparts.

Fig. 4.15 shows the relevant quantum capacitances and transconductances for the three devices. Fig. 4.15(a) shows C_{sq} for all three devices, as well as C_{dq} for the GFET, noting $C_{\mathrm{dq}} \approx 0$ for the CNFET and MOSFET; similarly, Fig. 4.15(b) shows g_{sq} for all three devices, as well as g_{dq} for the GFET, noting $g_{\mathrm{dq}} \approx 0$ for the CNFET and MOSFET.



(a)



(b)

Fig. 4.15. (a) Relevant quantum capacitances and (b) transconductances versus channel potential for a GFET, MOSFET, and CNFET. The curves are plotted for an applied gate bias of 0.2 V to 1 V.

As illustrated in Fig. 4.15(a), C_{sq} for the CNFET and MOSFET tend to flatten out with bias, which results in small values of the second-order coefficients (determined by the

first derivatives of the shown curves) in the Taylor-series expansion (4.1) for these devices; the coefficients in (4.2) are zero since $C_{\mathrm{dq}} \approx 0$ in a CNFET and MOSFET, due to the existence of a bandgap in the corresponding channel materials. On the other hand, *both* C_{sq} *and* C_{dq} show significant slope over most bias values for the GFET, causing the coefficients in *both* (4.1) *and* (4.2) to be pronounced for the GFET.

Similar results follow from inspection of Fig. 4.15(b), which suggests pronounced distortion from *both* (4.3) *and* (4.4) for the GFET, but only (4.3) for the MOSFET and neither for the CNFET.

Overall, the GFET will hence have second-order distortion contributions from *all four* nonlinear elements in Fig. 4.3, whereas only one or two of the components will play a role for the CNFET and MOSFET; the GFET thus exhibits the worst IIP2.

One subtle point about the GFET's IIP2 curve should be noted. Unlike the GFET's IIP3, its IIP2 peaks (sharply) at the gate bias for peak f_T. This outcome can be attributed to the minima in the drain components C_{dq} and g_{dq} at that bias point (as shown in Fig. 4.15 and earlier in Fig. 4.5), which makes the second-order coefficients determined by the first derivatives very small.

4.3.4. Extrinsic Linearity of GFET

a) Calculation of Parasitics

To calculate the extrinsic parasitics, the gate contact was assumed to be made of tungsten with dimensions $W \times L_g \times t_g$ of $1\ \mu\mathrm{m} \times 18\ \mathrm{nm} \times 60\ \mathrm{nm}$. Tungsten was chosen due to its closely matched work-function with graphene. From the resistivity of tungsten, the total resistance of the gate contact was calculated to be 220 Ω. The distributed gate resistance was then modeled as a lumped resistance, $R_{\mathrm{g,eff}} = 220/3\ \Omega$. The source and drain contact resistances were taken to be $R_s = R_d = 50\ \Omega$, near the theoretical minimum for graphene [90, 91]. The extrinsic capacitances were measured to be $C_{\mathrm{gd,ext}} = 40\ \mathrm{aF}$, $C_{\mathrm{gs,ext}} = 40\ \mathrm{aF}$, and $C_{\mathrm{sd,ext}} = 24\ \mathrm{aF}$ by simulating the open-pad structure in COMSOL [34].

TABLE 4.2. INTRINSIC AND EXTRINSIC CIRCUIT COMPONENTS OF THE GFET

Intrinsic Components	C_{ge} **[aF]**	1040
	C_{de} **[aF]**	233
	C_{se} **[aF]**	96
	C_{sq1} **[aF]**	784
	C_{dq1} **[aF]**	188
	g_{sq1} **[mS]**	27.74
	g_{dq1} **[mS]**	6.65
Extrinsic Components	$C_{gd,ext}$ **[aF]**	40
	$C_{gs,ext}$ **[aF]**	40
	$C_{sd,ext}$ **[aF]**	24
	$R_{g,eff}$ **[Ω]**	220/3
	R_s, R_d **[Ω]**	50

Table 4.2 lists the intrinsic and extrinsic circuit component values of the 1-μm wide GFET studied in this work; the bias-dependent values were calculated for gate and drain biases both equal to 0.5 V, and only the first-order coefficients are listed for the nonlinear elements.

b) Extrinsic Linearity Features of a GFET

Once developed, the final extrinsic equivalent circuit was simulated in MWO [30] and the resulting IIP3 values are plotted versus gate bias in Fig. 4.16. For the 1-µm wide device, the external parasitics were found to slightly degrade the device linearity, but the signature shape identified from the intrinsic device remains, as demonstrated in Fig. 4.16.

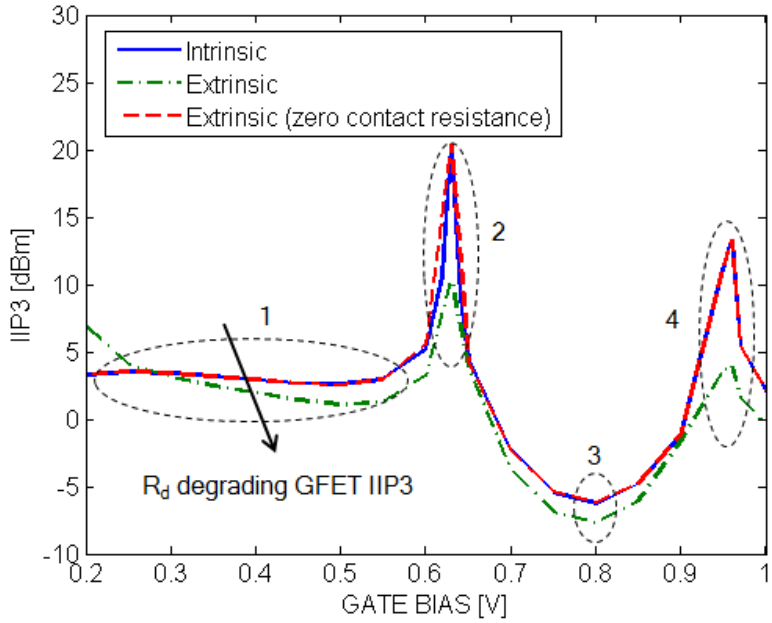


Fig. 4.16. Intrinsic and extrinsic IIP3 versus gate bias.

By selectively removing the parasitics one by one from the circuit in Fig. 4.3 and solving in MWO, we found that the extrinsic *capacitances* do not affect the RF linearity of GFETs; rather, it is the *contact resistances*. The following discussion identifies the contribution of the contact resistances to GFET linearity.

c) Impact of Drain Contact Resistance

Our investigation showed that the drain contact resistance is primarily responsible for degrading the overall RF linearity of the GFET. The potential drop across the drain contact resistance added to the output voltage $v_{\mathrm{d,ext}}$ results in a larger intrinsic drain voltage v_d in the circuit of Fig. 4.3. As discussed in Section 4.3.2(a), a larger v_d increases the distortion from the nonlinear drain components in Fig. 4.3, a phenomenon unique to GFETs, and makes the device more nonlinear, by increasing the control voltage $(v_d - v_{\mathrm{scf}})$ in (4.2) and (4.4).

It should be mentioned that the source contact resistance will tend to improve the linearity of the device slightly due to its well-known feedback effect in the common-

source configuration [p.101, 43], but any such improvement is dominated by the degrading effect of the drain resistance.

The gate contact resistance is small enough in the 1-µm wide device that it does not affect the linearity; we will shortly consider a wider device to isolate its effect.

Simulating the extrinsic circuit with zero R_d while retaining the parasitic capacitances results in identical linearity between the extrinsic and intrinsic devices, as shown in Fig. 4.16. A small drain resistance is hence essential to making a GFET as linear as possible.

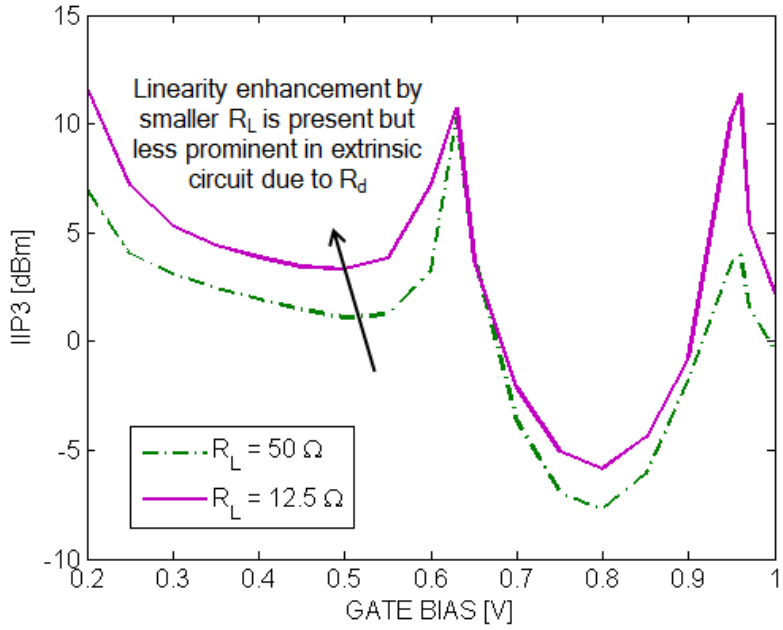


Fig. 4.17. Effect of R_L on GFET extrinsic IIP3. The improvement in IIP3 with a reduction in R_L is less pronounced than in the intrinsic case shown in Fig. 4.12.

One other note should be made about the impact of the drain resistance. In Section 4.3.2(b), it was shown that a small R_L has the potential to improve the GFET linearity by reducing the swing of the drain voltage v_d (Fig. 4.12). However, the presence of the drain contact resistance makes it impossible to lower the swing of v_d enough to improve linearity significantly. Fig. 4.17 shows the effect of variation in R_L on linearity for the extrinsic GFET. The reduction in R_L that improved the linearity by almost 10 dB in the intrinsic circuit only improves the linearity by 2.2 dB in the extrinsic circuit. Keeping the

drain contact resistance low is hence also important to allow for potential linearity improvement by adjusting R_L.

d) Effect of Gate Contact Resistance

To examine the impact of $R_{g,eff}$, a wider device ($W \geq 10$ µm) must be considered, where $R_{g,eff}$ is appreciable. Fig. 4.18 shows that the linearity of the 10-µm wide device improves significantly when the effects of the external parasitics are included. The drain contact resistance still degrades the linearity, but the degradation is canceled by an even greater improvement in linearity due to the gate contact resistance $R_{g,eff}$. As device width increases, the gate contact resistance can hence improve linearity, but of course this would come at the expense of reduced power gain.

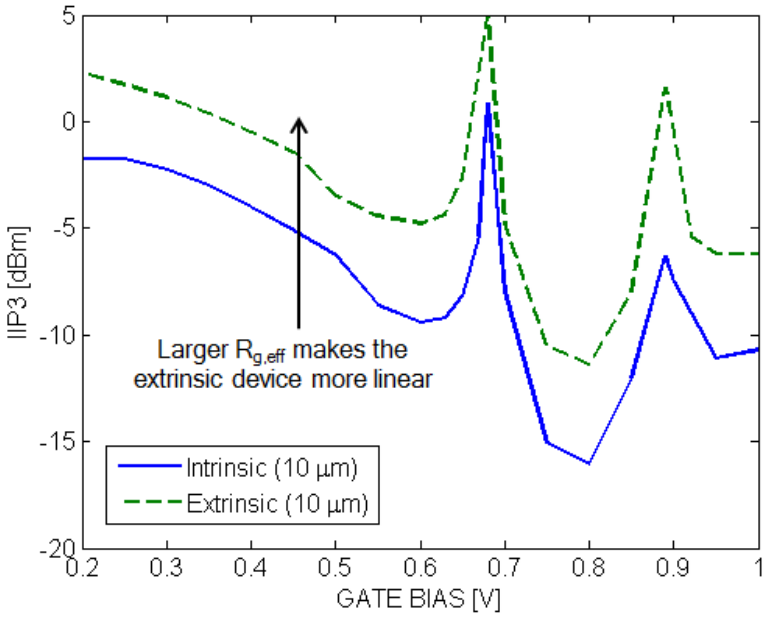
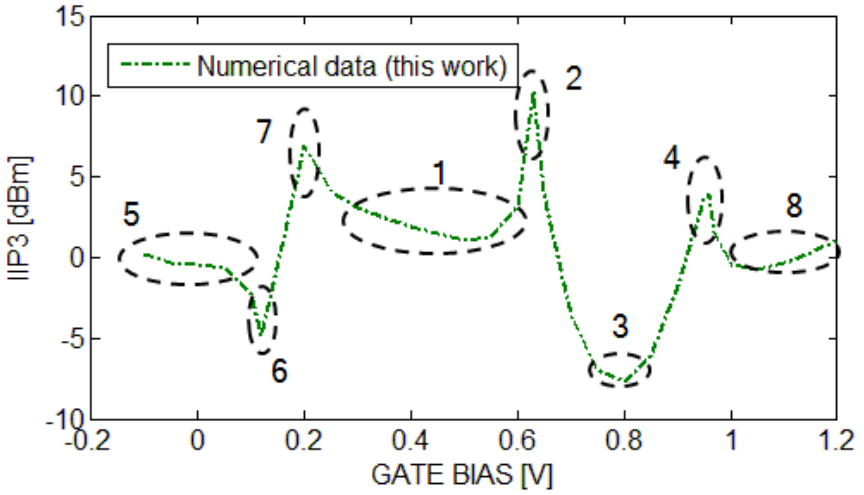
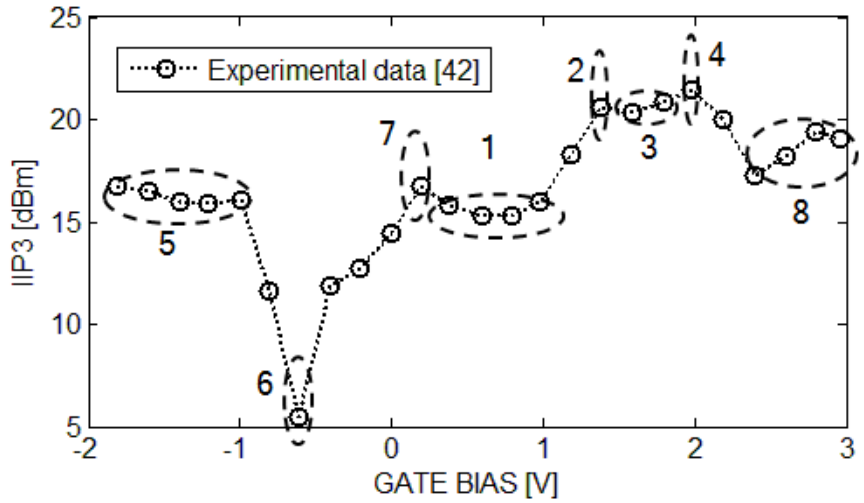


Fig. 4.18. Intrinsic and extrinsic IIP3 versus gate bias for a 10-µm wide GFET.

4.4. Qualitative Comparison with Experimental Results



(a)



(b)

Fig. 4.19: Qualitative comparison of (a) numerical (extrinsic) IIP3 values of the GFET under investigation with (b) experimental data [42].

Finally, we compare our IIP3 values with experimental results. As discussed in Section 4.1, the experimental studies of the RF linearity of GFETs have all examined long-channel devices ($L \geq 250$ nm), which makes a *direct* comparison with our simulation results impossible. However, a *qualitative* comparison with the reported IIP3

values in [42] (Fig. 4.19) shows that the key signature of the GFET IIP3 (regions 1 to 4), as identified in Section 4.3.1, is present even in a long-channel ($L = 700$ nm) fabricated device.

Fig. 4.19 shows that the peaks at points 2 and 4 are diminished in the experimental IIP3 and that the dip around point 3 is also less prominent. The differences between the experimental data and the numerical results are most likely due to the nonidealities in the practical device that our model does not consider, such as scattering. The fabricated GFET in [42] is a long-channel device with a channel length of 700 nm. The transport in this device is therefore subject to scattering, which is neglected in the short-channel GFET ($L = 18$ nm) considered in this study. As discussed in Section 4.3.1, the formation of the peaks at points 2 and 4 are dependent on the phase relationship of the distortion generated by the second- and third-order coefficients of the drain components C_{dq} and g_{dq}. The presence of scattering in the long-channel device can be expected to change this phase relationship and thereby diminish the peaks.

Extending the gate bias values beyond the 0.2 V to 1 V range used throughout our study thus far shows that the developed model is capable of capturing *all* the features present in the experimental IIP3 curve [42].[10] Regions 5, 6, 7, and 8 in the extended plot in Fig. 4.19(a) clearly mirror the corresponding regions in Fig. 4.19(b). The mechanism behind these regions can be revealed by examining our developed nonlinear model.

- Regions 5 and 8 are very similar to region 1 and the IIP3 values in these regions are relatively insensitive to gate bias. As in region 1, the source and drain quantum capacitances and transconductances vary linearly with voltage in these regions and therefore the IIP3 remains almost constant.
- Point 6 occurs at the point of minimum conduction, such that the small-signal transconductance g_m is zero. This zero transconductance results in a small-signal gain A_v of zero at this particular bias point. The output of the device

[10] In comparing the numerical and experimental data in parts (a) and (b) of Fig. 4.19, the *actual* gate bias and IIP3 values are not important; these will not overlap, as the two devices involved have different channel lengths. Of relevance is the *relative* positions of the identified regions and points with respect to gate bias, and the resulting *signature* in the IIP3 behavior.

therefore contains distortion due to the nonlinear circuit components, but the fundamental frequency component is absent. This makes the device extremely nonlinear at this bias point, and therefore the IIP3 exhibits a sharp drop.

- Region 7 has an IIP3 sweet-spot. By separately examining the nonlinearity of the source and drain components in the small-signal circuit, we found that the distortion from both the source and drain components at point 7 are of equal magnitude and opposite phase. This results in a destructive combination of distortion components, making the device very linear.

For the purpose of this thesis, the similarity of the curves in the two parts of Fig. 4.19 provides a striking qualitative validation of our modeling approach and our resulting observations on the linearity of GFETs, as reported in our work [17] submitted to the *IEEE Transactions on Microwave Theory and Techniques*. However, the similarity of the curves in Fig. 4.19 begs a more detailed investigation and discussion that clarifies why the signature behavior of linearity, which is apparently present at all channel lengths, is governed (for graphene) by features of a ballistic transport model. This investigation and discussion will be reported in work to be published separately; here, the most important outcome is that the similarity of experiment and our simulations provides a strong validation of our approach and conclusions.

4.5. Conclusions

The following conclusions can be drawn regarding the RF linearity potential of GFETs.

1. The IIP3 versus gate bias curve of the GFET has four distinct features (Fig. 4.4). A constant linearity region, two sharp peaks, and a large dip.
2. The linear DOS of graphene results in a linear quantum capacitance and transconductance versus voltage relationship in GFETs (Fig. 4.5) at low gate bias (0.2 V$\leq v_G \leq$ 0.5 V), which is responsible for the constant linearity region.

3. Depending on the gate bias, the GFET linearity is dictated either by distortion generated by second-order coefficients or by third-order coefficients in the Taylor-series expansions (4.1)-(4.4) of the nonlinear components. A destructive combination of distortions from the two mechanisms in the transition regions create sharp peaks in the IIP3 curve (Fig. 4.6).
4. The GFET offers its worst linearity at peak f_T (Fig. 4.4).
5. Over all gate bias values, the distortion generated in the nonlinear drain components dictate the GFET linearity (Fig. 4.9). This is an outcome of the zero bandgap of graphene. It also makes the RF linearity highly sensitive to variations in drain bias (Fig. 4.10) and potentially load resistance (Fig. 4.12).
6. In terms of third-order distortion, the GFET's performance is comparable to its MOSFET and CNFET counterparts (Fig. 4.13), with the distinguishing feature that the peak IIP3 does not occur at peak f_T.
7. Due to its linear DOS and lack of a bandgap, the second-order distortion is much worse in a GFET than in its competitors (Fig. 4.14).
8. The extrinsic IIP3 retains the key features (signatures) of the intrinsic IIP3 (Fig. 4.16).
9. Parasitic capacitances have a minimal impact on GFET linearity.
10. The drain contact resistance degrades the linearity of a GFET (Figs. 4.16, 4.17), while the source resistance has minimal impact; this occurs due to the drain dominance of GFET linearity (conclusion 5). In wide devices ($W \geq 10$ μm), the gate contact resistance can make the device more linear (Fig. 4.18), but of course it will also degrade the power gain.

Overall, the most important outcomes of this work are the identification of the signature behavior (Figs. 4.4 and 4.16) and the drain dependence of graphene linearity (shown, for example, in Fig. 4.10). We also showed that graphene has the potential to offer third-order linearity at least comparable to CNFETs and MOSFETs (Fig. 4.13), but suffers from worse second-order linearity (Fig. 4.14). The load-resistance dependency creates a unique opportunity to improve the linearity in GFETs by using smaller loads,

but at the cost of reduced voltage and power gain (Table 4.1). All these key outcomes are intimately tied to the lack of a bandgap and linear DOS of graphene.

Chapter 5

Conclusions and Future Work

5.1. Summary of Contributions

For convenience, we will summarize here the main contributions from each stage of work. Further details can be found in the preceding chapters, where each stage is discussed in turn.

5.1.1. Stage 1 (Chapter 2)

The contributions from the first stage, entitled, "RF Linearity Potential of Carbon-Nanotube Transistors Versus MOSFETs," are as follows:

a. The RF linearity of CNFETs is shown to be comparable to their MOSFET counterparts, contrary to earlier predictions [22].

b. The nonlinear quantum capacitance is shown to be a major source of nonlinearity in CNFETs, unlike conventional MOSFETs.

c. The tube pitch is shown to affect the CNFET linearity for sufficiently wide devices, with the linearity improving in devices having a smaller pitch.

Collectively, the most important outcomes of the work are a physical understanding of the nonlinear mechanism in CNFETs under ballistic conditions and how their performance compares to their MOSFET counterparts in terms of RF linearity.

5.1.2. Stage 2 (Chapter 3)

The contributions from the second stage, entitled, "A Modified Top-of-the-Barrier Solver for Graphene and Its Application to Predict RF Linearity," are as follows:

a. A novel, nonlinear, small-signal circuit model for GFETs that includes the effects of varying densities of states and band-to-band tunneling is developed.

b. The RF linearity of GFETs is shown to be highly sensitive to drain biasing, unlike their CNFET and MOSFET counterparts.

c. The RF linearity of GFETs is shown to be comparable to their CNFET and MOSFET counterparts under proper drain biasing.

Collectively, the most important outcomes of the work are a modified top-of-the-barrier approach suitable for studying graphene linearity, *i.e.*, one that includes varying densities of states and band-to-band-tunneling, and a preliminary indication of how the linearity of these devices compares to their CNFET and MOSFET counterparts.

5.1.3. Stage 3 (Chapter 4)

The contributions from the third stage, entitled, "RF Linearity Performance Potential of Short-Channel Graphene Field-Effect Transistors," are as follows:

a. The IIP3 versus gate bias curve of the GFET is shown to have four distinct features; a constant linearity region, two sharp peaks, and a large dip.

b. Over all gate bias values, the distortion generated in the nonlinear drain components is shown to dictate the GFET linearity. This is an outcome of the zero bandgap of graphene. It also makes the RF linearity highly sensitive to variations in drain bias and potentially load resistance.

c. In terms of third-order distortion, the RF linearity of GFETs is shown to be comparable to their MOSFET and CNFET counterparts.

d. The extrinsic IIP3 is shown to retain the key features (signature) of the intrinsic IIP3; however, the drain contact resistance is shown to slightly degrade the overall linearity of a GFET.

Collectively, the most important outcomes of the work are the identification of the signature behavior and the drain dependence of graphene linearity and how the linearity of GFETs compares to their MOSFET and CNFET counterparts.

5.2. Future Projects

In this work, we have investigated the RF linearity of carbon-based nanoscale transistors. The methodology, however, is transferable to other technologies. The developed nonlinear small-signal equivalent circuit can be used to describe any ballistic nanoscale transistor and hence can be used to assess their high frequency figures of merit. One natural extension of this Ph.D. work can, therefore, be to probe the RF linearity of other emerging nanoscale transistors.

Throughout this work, we have assumed ballistic transport to keep things simple while still providing useful insights into the performance potential of these emerging carbon-based device technologies. However, in order to completely model the behavior of a practical FET, one has to consider the presence of elastic and inelastic scattering; such scattering is more important in long-channel prototype devices, where the channel lengths can be hundreds of nanometers or even microns, but it can also play a role even when the channel lengths are only a few nanometers, where the transport is often described as "quasi"-ballistic as opposed to "purely" ballistic. Another extension of this Ph.D. work can, therefore, be to study the impact of scattering on the RF linearity of these emerging nanoscale devices.

Two associated projects for future consideration are described below.

5.2.1. Project A: Effect of Bandgap on RF Linearity of MoS_2 Transistors Versus GFETs

a) Motivation for Project A

The unique physical and electrical properties of 2D graphene have started a massive interest in the use of 2D materials as a channel material in field-effect transistors. Among the 2D materials being considered, MoS_2 has made its mark due to the presence of a 1.8

eV direct bandgap [92, 93]. The wide bandgap of MoS_2 has made it a particularly promising material in the future of digital electronics [94-96], and the high mobility [95] of these transistors has created interest in their use in analog electronics [97-99]. Recent studies have emphasized that the lack of a bandgap in graphene not only makes it difficult to use the material in digital electronics but also degrades its high frequency performance [10, 27, 100]. Being free from this shortcoming, MoS_2 might hence hold promise in RF electronics in addition to digital applications. Compact nonlinear modeling of MoS_2 FETs will enable a quick assessment of their high-frequency capabilities, such as RF linearity. In this part of the work, we will develop a nonlinear small-signal circuit for the MoS_2 FET based on the top-of-the-barrier approach. Simulation and analysis of the circuit will reveal whether the wide bandgap of MoS_2 results in attractive RF linearity along with attractive digital properties, such as a large on-off current ratio (I_{on}/I_{off}) [94-96].

b) Methodology for Project A

The nonlinear small-signal circuit for a MoS_2 FET will be developed based on the "top-of-the-barrier" approach [13]. Existing literature on the band structure of MoS_2 indicates a parabolic nature of the band profile at low energies [101]. Using the effective-mass approach, the best fitted parabolic band structure will be used inside the top-of-the-barrier model. Once developed, the nonlinear solver in Microwave Office [30] will be used to analyze the circuit and the linearity figures of merit will be generated. The effect of device parameters such as channel width and oxide thickness will be investigated.

c) Anticipated Significance and Outcome of Project A

One outcome of this part of the work will be a compact nonlinear circuit model for MoS_2 FETs. In addition, a comparison of RF linearity with GFETs will reveal whether the bandgap of MoS_2 plays a role in linearizing the device behavior. The developed

model will be used to study the effects of device parameters on RF linearity and will hence further classify the RF capabilities of MoS_2 as a channel material.

5.2.2. Project B: Compact Nonlinear Modeling of Scattering in GFETs and Its Effect on RF Linearity

In order to develop a compact, nonlinear, small-signal circuit model for GFETs that includes all relevant physics, we will introduce the effects of scattering to the ballistic model of Chapter 3. The improved model will be calibrated with a sophisticated and numerically involved NEGF solver for GFETs that is currently being developed by my colleague, Kyle D. Holland. The compact model will enable us to assess the effects of scattering on device performance of GFETs under different conditions in a time-efficient way. Nonlinear analysis of the circuit will reveal the linearity of a practical GFET and will shed insight into the effects of scattering on RF linearity.

Appendix A: Top-of-the-Barrier Model

Overview

The top-of-the-barrier model provides a simple method to model the charge and current in any *ballistic* (collisionless) transistor. First developed by Natori [12] in 1994, the model is based on the concept that in a ballistic MOSFET, the so-called "top of the barrier," defined as the point where the conduction band reaches its highest energy, creates a *bottleneck* for transport between the source and the drain. In this appendix, we briefly discuss the key ideas behind the top-of-the-barrier model [12-14].

Visualization

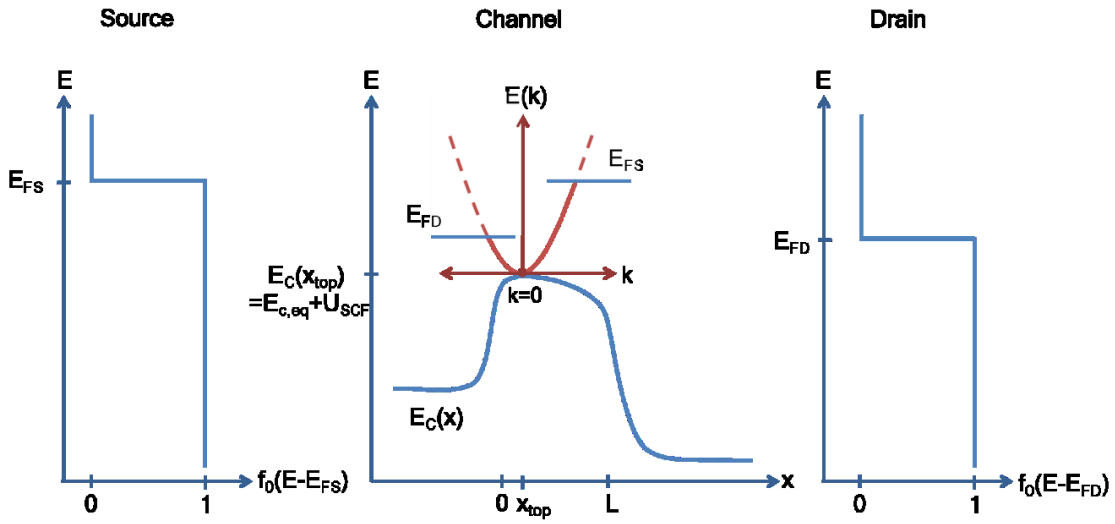


Fig. A.1. An illustration of the conduction-band edge $E_C(x)$ versus transport direction x in a MOSFET. The E–k relation at $x = x_{\text{top}}$, the location of the top of the source–to-drain energy barrier, is also shown. At 0 K, the positive- and negative-going k-states get filled up to the source and drain Fermi levels, respectively, as indicated by the solid portions of the dispersion curve.

The basic situation is sketched in Fig. A.1. Shown is the conduction-band edge E_C versus position x (along the transport direction) in a one-dimensional ballistic MOSFET. We have also shown the source and drain to the left and right of the sketch, respectively, both of which are modeled as ideal reservoirs, governed by their local distribution functions $f_0(E - E_{\rm FS})$ and $f_0(E - E_{\rm FD})$, where $f_0(E - \theta)$ refers to the equilibrium Fermi-Dirac form,

$$f_0(E-\theta) = \frac{1}{1 + e^{(E-\theta)/k_B T}} \tag{A1}$$

with k_B being Boltzmann's constant and T being the temperature, and where $E_{\rm FS} = \mu_0 - qV_S$ and $E_{\rm FD} = \mu_0 - qV_D$ are the local Fermi levels in the source and drain, respectively, with μ_0 being the equilibrium Fermi level and q being the magnitude of the electronic charge. At the location $x = x_{\rm top}$, we have superimposed the local energy-versus-wave-vector dispersion relation[11] (E versus k), and we have marked the source and drain Fermi levels for reference. The value of the conduction-band energy at the top of the barrier is also marked, and can be written as

$$E_C\left(x_{\rm top}\right) = E_{C,\rm eq} + U_{\rm SCF} \tag{A2}$$

where $E_{C,\rm eq}$ is the equilibrium position of the conduction band at the top of the barrier and $U_{\rm SCF}$ is the amount by which it is shifted by the local electrostatic potential.

Transport Considerations

The top-of-the-barrier approach can now be understood by first focusing on the electron occupancy of the available states on the E-k dispersion curve at $x = x_{\rm top}$. Under *purely ballistic* (absolutely collisionless) conditions, states on the right half of the curve (*i.e.*, for $k > 0$) must be occupied by electrons that came from the source, since the

[11] For visualization purposes, we assume a parabolic dispersion relation in the sketch; the actual dispersion will be determined by the channel material.

velocity of such electrons is positive[12]: $v_k = (1/\hbar)(dE/dk) > 0$ for $k > 0$, where $\hbar = h/(2\pi)$ and h is Planck's constant. Similarly, states on the left half of the curve (*i.e.*, for $k < 0$) must be occupied by electrons that came from the drain, since the velocity of such electrons is negative: $v_k = (1/\hbar)(dE/dk) < 0$ for $k < 0$. Moreover, under ballistic transport, the distributions of the electrons in each half of the dispersion curve must be governed by the distributions of their originating reservoirs; this follows because collisions are absent and hence there is no mechanism to alter the originating distributions. It is thus possible to write the following expressions for the numbers of electrons n^+ and n^- in the two halves of the dispersion curve [13]:

$$n^+ = 2\sum_{k>0} f_0[E(k) - E_{\mathrm{FS}}] = \int_{E_c(x_{\mathrm{top}})}^{\infty} D(E) f_0(E - E_{\mathrm{FS}}) dE \tag{A3}$$

$$n^- = 2\sum_{k<0} f_0[E(k) - E_{\mathrm{FD}}] = \int_{E_c(x_{\mathrm{top}})}^{\infty} D(E) f_0(E - E_{\mathrm{FD}}) dE \tag{A4}$$

where the factor of two preceding the sum accounts for spin degeneracy, the sum over k can be converted to an integral over E in the usual way [102] and $D(E)$ is the density of states at the top of the barrier, which can be found from knowledge of the E-k dispersion.

Equations (A3) and (A4) are the so-called *transport equations* for charge at the top of the barrier. They can be evaluated if the top-of-the-barrier energy $E_c(x_{\mathrm{top}})$ in (A2) is known, and this in turn requires knowledge of the electrostatic potential U_{SCF}.

[12] Since collisions are absent, there is no mechanism for an electron injected from the drain to have a positive velocity.

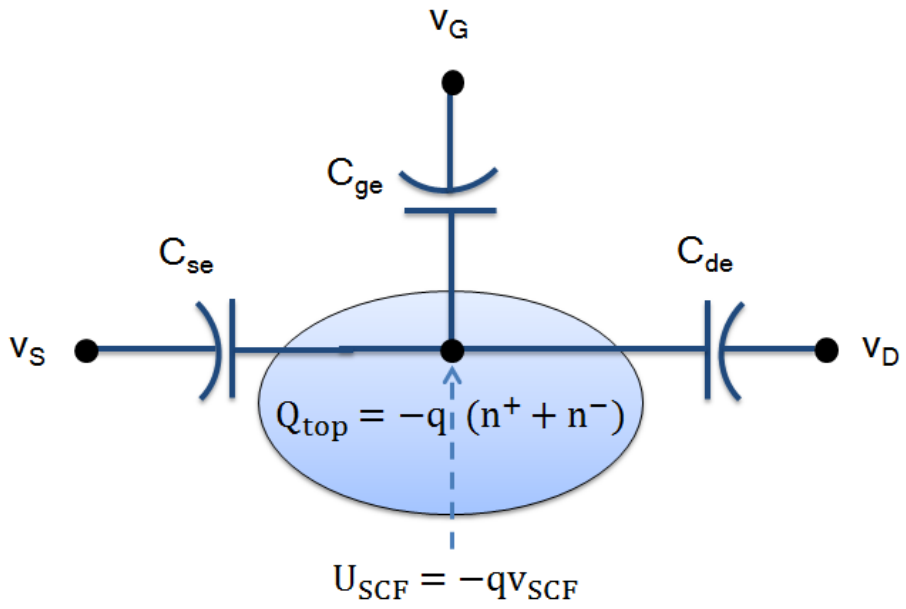


Fig. A.2. Simple circuit model for the electrostatic behavior of a ballistic MOSFET. The electrostatic potential at the top of the barrier U_{SCF} is controlled by the voltages applied to the gate, drain, and source through the three illustrated capacitors.

Electrostatic Considerations

The top-of-the-barrier model assumes that the potential U_{SCF} can be found by way of the simplified circuit model displayed in Fig. A.2. This model assumes that the external voltages impact the potential at the top of the barrier (the central node in the circuit) by way of three electrostatic capacitors, labeled here as C_{se}, C_{ge}, and C_{de}, and whose values must be known.

The net charge on the inner capacitor plates can be written as

$$Q_{\mathrm{top}} + qn_0 = C_{\mathrm{se}}(v_{\mathrm{SCF}} - v_S) + C_{\mathrm{ge}}(v_{\mathrm{SCF}} - v_G) + C_{\mathrm{de}}(v_{\mathrm{SCF}} - v_D) \qquad \text{(A5)}$$

where qn_0 is the equilibrium channel charge and v_{SCF} is the local electrostatic potential. Rearranging (A5) with $U_{\mathrm{SCF}} \equiv -qv_{\mathrm{SCF}}$, one can easily obtain the result

$$U_{\mathrm{SCF}} = -q(\alpha_G v_G + \alpha_D v_D + \alpha_S v_S) + q^2(n - n_0)/C_\Sigma \qquad \text{(A6)}$$

where $\alpha_G \equiv C_{\rm ge}/C_\Sigma$, $\alpha_D \equiv C_{\rm de}/C_\Sigma$, and $\alpha_S \equiv C_{\rm se}/C_\Sigma$ are capacitance ratios that characterize the device and $C_\Sigma = C_{\rm ge} + C_{\rm de} + C_{\rm se}$ is the total electrostatic capacitance. Typical values of the capacitance ratios in a conventional silicon MOSFET are α_G =0.87, α_D =0.033, and α_S =0.017 [13].

Self-Consistent Solution

Equations (A3)--(A4) [with $E_C(x_{\rm top})$ defined by (A2)] and (A6) represent a set of coupled nonlinear equations that can be solved for the two unknowns n and $U_{\rm SCF}$.[13] Since a solution of the equations involves *self-consistency* between transport and electrostatic considerations, the resulting potential $U_{\rm SCF}$ is aptly dubbed the "self-consistent potential."

Once $U_{\rm SCF}$ is known, the drain-to-source current can be calculated by summing the contribution from each occupied state of the E-k dispersion at $x = x_{\rm top}$. For the positive-going electrons,

$$i^+ = -2q \sum_{k>0} v_k \frac{f_0[E(k) - E_{\rm FS}]}{L} \tag{A7}$$

where $(f_0\,[E(k) - E_{FS}])/L$ represents the (normalized) electron density in a 1D channel (such as carbon nanotube) due to occupancy of the state k, with L being the channel length, and where the preceding factor of 2 accounts for spin degeneracy. The sum in (A7) can again be converted to an integral over energy in the usual way [102] to obtain

$$i^+ = -\frac{2q}{h} \int_{E_c(x_{\rm top})}^{\infty} f_0(E - E_{\rm FS})\, dE \tag{A8}$$

Similarly, for the negative-going electrons, we have

[13] The remaining parameters appearing in these equations are presumed to be known for the MOSFET being modeled.

$$i^- = -2q \sum_{k<0} v_k \frac{f_0[E(k) - E_{\mathrm{FD}}]}{L} = -\frac{2q}{h} \int_{E_c(x_{\mathrm{top}})}^{\infty} f_0(E - E_{\mathrm{FD}}) dE \tag{A9}$$

Finally, the drain-to-source current can be written as

$$i_{\mathrm{DS}} = i^+ - i^- = -\frac{2q}{h} \int_{E_c(x_{\mathrm{top}})}^{\infty} [f_0(E - E_{\mathrm{FS}}) - f_0(E - E_{\mathrm{FD}})]\, dE \tag{A10}$$

Summary

Based on the above visualization and formalism, the key equations of the top-of-the-barrier model have been boxed. Input parameters to the model are the electrostatic capacitances in Fig. A.2, the equilibrium Fermi level μ_0, and the E-k dispersion, or equivalently, the density of states $D(E)$ of the channel material. With these parameters specified for a device of interest, the idea is to solve (A3) and (A4) self-consistently with (A6) for U_{SCF}, where U_{SCF} appears through $E_c(x_{\mathrm{top}})$ according to (A2). Once U_{SCF} and hence $E_c(x_{\mathrm{top}})$ are known, the current can then be determined from (A10).

Appendix B: Derivation of the Condition for "Inherent Linearity"

The condition for inherent linearity was first derived by Baumgardner *et al.* in [22]. Here, we outline the derivation to be consistent with our top-of-the-barrier [12-14] equations and for the convenience of the readers.

Under the assumption of a sufficiently high drain bias to neglect the drain-injected charge at the top of the source-drain barrier, the transport equations for the channel charge and current can be written as follows [12-14]:

$$\lambda = -\frac{q}{2}\int_{E_{\mathrm{C,eq}}}^{\infty} D(E) f[E - qv_{\mathrm{SCF}} - E_F + qv_S]dE \tag{B1}$$

and

$$i_T \approx i_{\mathrm{TS}} = \frac{4q}{h}\int_{E_{\mathrm{C,eq}}}^{\infty} f[E - qv_{\mathrm{SCF}} - E_F + qv_S]dE \tag{B2}$$

where $E_{\mathrm{C,eq}}$ is the equilibrium position of the conduction band and v_{SCF} is defined to be zero at equilibrium. In addition, neglecting the capacitances C_{se} and C_{de} for simplicity, the electrostatics demands the following relation for the channel charge [12-14]:

$$\lambda = -C_{\mathrm{ge}}(v_G - v_{\mathrm{SCF}}). \tag{B3}$$

Equating (B1) and (B3), and setting v_S to zero, we get

$$-C_{\mathrm{ge}}(v_G - v_{\mathrm{SCF}}) = -\frac{q}{2}\int_{\mathrm{E_{C,eq}}}^{\infty} D(E) f[E - qv_{\mathrm{SCF}} - E_F]dE. \tag{B4}$$

Solving (B4) for v_{SCF} then results in

$$v_{\mathrm{SCF}} = v_G + \frac{-\frac{q}{2}\int_{E_{\mathrm{C,eq}}}^{\infty} D(E) f[E - qv_{\mathrm{SCF}} - E_F]dE}{C_{\mathrm{ge}}} \equiv v_G + \frac{\lambda}{C_{\mathrm{ge}}}. \tag{B5}$$

For the case $\left|\frac{\lambda}{C_{\mathrm{ge}}}\right| \ll v_G$, the right side of (B5) collapses to v_G, and the channel potential follows the gate bias directly, *i.e.*, $v_{\mathrm{SCF}} \approx v_G$. Using this solution for v_{SCF} with $v_S \equiv 0$, the transport current can then be rewritten from (B2) as

$$i_T = \frac{4q}{h}\int_{E_{\mathrm{C,eq}}}^{\infty} f[E - qv_G - E_F]dE$$

$$= k_B T \frac{4q}{h} \ln\left[1 + e^{\frac{q(v_G - v_T)}{k_B T}}\right] \tag{B6}$$

where $qv_T = E_{\mathrm{C,eq}} - E_F$ is the height of the source-drain barrier at equilibrium. The value of qv_T is typically small for nanotubes; for example, for the CNFET considered in our study, $qv_T = 0.25$ eV.

Once $\frac{q(v_G - v_T)}{k_B T} \gg 1$, the right side of (B6) becomes the logarithm of an exponent, and the transport current becomes "inherently linear" with v_G:

$$i_T = \frac{4q^2}{h}(v_G - v_T). \tag{B7}$$

References

[1] C. A. Mack, "Fifty years of Moore's law," *IEEE Trans. Semicond. Manuf.,* vol. 24, pp. 202-207, May 2011.

[2] M. Bohr, "Moore's law in the innovation era," in *Design for Manufacturability through Design-Process Integration V,* Mar. 2011, pp. 797402-1-8.

[3] S. E. Thompson, "Power, cost and circuit IP reuse: The real limiter to Moore's law over the next 10 years," in *2010 International Symposium on VLSI Technology, Systems, and Applications (VLSI-TSA),* Apr. 2010, pp. 88-89.

[4] K. J. Kuhn, "Moore's law past 32 nm: Future challenges in device scaling," in *2009 13th International Workshop on Computational Electronics (IWCE),* Aug. 2009, pp. 1-6.

[5] T. Dürkop, S. A. Getty, E. Cobas, and M. S. Fuhrer, "Extraordinary mobility in semiconducting carbon nanotubes," *Nano Lett.*, vol. 4, no. 1, pp. 35-39, Jan. 2004.

[6] K. I. Bolotin, K. J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. L. Stormer, "Ultrahigh electron mobility in suspended graphene," *Solid State Comm.,* vol. 146, pp. 351-355, Jun. 2008.

[7] E. Pop, V. Varshney, and A. K. Roy, "Thermal properties of graphene: Fundamentals and applications," *MRS Bull.*, vol. 37, pp. 1273-1281, Dec. 2012.

[8] K. Nishiguchi and K. Ocha, "Conductance quantization in nanoscale vertical structure silicon field-effect transistors with a wrap gate," *Appl. Phys. Lett.*, vol. 76, no. 20, pp. 2922-2924, May 2000.

[9] "Radio frequency and analog/mixed-signal technologies for wireless communications," *International Technology Roadmap for Semiconductors.* (2012) [Online]. Available: http://www.itrs.net/Links/2012ITRS/ 2012Tables/PIDS_2012 Tables.xlsx.

[10] K. D. Holland, N. Paydavosi, N. Neophytou, D. Kienle, and M. Vaidyanathan, "RF

performance limits and operating physics arising from the lack of a bandgap in graphene transistors," *IEEE Trans. Nanotechnol.,* vol. 12, no. 4, pp. 566-577, Jul. 2013.

[11] S. Ahmed, N. Paydavosi, A. U. Alam, K. D. Holland, D. Kienle, and M. Vaidyanathan, "Impact of the substrate material on the RF performance of carbon-nanotube transistors," *IEEE Trans. Nanotechnol.,* vol. 13, no. 1, pp. 123-135, Jan. 2014.

[12] K. Natori, "Ballistic metal-oxide-semiconductor field effect transistor," *J. Appl. Phys.,* vol. 76, pp. 4879-4890, Oct. 1994.

[13] A. Rahman, J. Guo, S. Datta, and M. S. Lundstrom, "Theory of ballistic nanotransistors," *IEEE Trans. Electron Devices,* vol. 50, no. 9, pp. 1853-1864, Sep. 2003.

[14] S. Hasan, S. Salahuddin, M. Vaidyanathan, and M. A. Alam, "High-frequency performance projections for ballistic carbon-nanotube transistors," *IEEE Trans. Nanotechnol.,* vol. 5, no. 6, pp. 14-22, Jan. 2006.

[15] A. U. Alam, C. M. S. Rogers, N. Paydavosi, K. D. Holland, S. Ahmed, and M. Vaidyanathan, "RF linearity potential of carbon-nanotube transistors versus MOSFETs," *IEEE Trans. Nanotechnol.,* vol. 12, no. 3, pp. 340-351, May 2013.

[16] A. U. Alam, K. D. Holland, S. Ahmed, D. Kienle, and M. Vaidyanathan, "A modified top-of-the-barrier model for graphene and its application to predict RF linearity," in *International Conference on Simulation of Semiconductor Processes and Devices, 2013 (SISPAD),* Sep. 2013, pp. 155-158.

[17] A. U. Alam, K. D. Holland, M. Wong, S. Ahmed, D. Kienle, P. Gudem, and M. Vaidyanathan, “RF linearity performance potential of short-channel graphene field-effect transistors,” Submitted to *IEEE Trans. Microwave Theory Tech.,* Manuscript ID TMTT-2014-08-0894, 12 pages, Aug. 2014.

[18] C. Rutherglen, D. Jain, and P. Burke, "Nanotube electronics for radiofrequency applications," *Nat. Nanotechnol.,* vol. 4, pp. 811-819, Dec. 2009.

[19] P. J. Burke, "AC performance of nanoelectronics: Towards a ballistic THz nanotube transistor," in *2003 International Semiconductor Device Research Symposium (ISDRS),* Dec. 2003, pp. 1981-1986.

[20] L. C. Castro and D. L. Pulfrey, "Extrapolated f_{max} for carbon nanotube field-effect transistors," *Nanotechnology,* vol. 17, pp. 300-304, Jan. 2006.

[21] N. Paydavosi, J. P. Rebstock, K. D. Holland, S. Ahmed, A. U. Alam, and M. Vaidyanathan, "RF performance potential of array-based carbon-nanotube transistors-part II: Extrinsic results," *IEEE Trans. Electron Devices,* vol. 58, pp. 1941-1951, Jul. 2011.

[22] J. E. Baumgardner, A. A. Pesetski, J. M. Murduck, J. X. Przybysz, J. D. Adam, and H. Zhang, "Inherent linearity in carbon nanotube field-effect transistors," *Appl. Phys. Lett.,* vol. 91, pp. 052107-1-3, Jul., 2007.

[23] A. Curutchet, D. Theron, M. Werquin, D. Ducatteau, H. Happy, G. Dambrine, J. M. Bethoux, V. Derycke, and C. Gaquiere, "Nonlinear characterization and modeling of carbon nanotube field-effect transistors," *IEEE Trans. Microwave Theory Tech.,* vol. 56, pp. 1505-1510, Jul. 2008.

[24] C. Wang, A. Badmaev, A. Jooyaie, M. Bao, K. L. Wang, K. Galatsis, and C. Zhou, "Radio frequency and linearity performance of transistors using high-purity semiconducting carbon nanotubes," *ACS Nano,* vol. 5, pp. 4169-4176, May 2011.

[25] F. Schwierz, "Graphene transistors," *Nat. Nanotechnol.,* vol. 5, pp. 487-496, May 2010.

[26] L. Liao, J. Bai, R. Cheng, Y. Lin, S. Jiang, Y. Qu, Y. Huang, and X. Duan, "Sub-100 nm channel length graphene transistors," *Nano Letters,* vol. 10, pp. 3952-3956, Sep. 2010.

[27] S. O. Koswatta, A. Valdes-Garcia, M. B. Steiner, Yu-Ming Lin, and P. Avouris, "Ultimate RF performance potential of carbon electronics," *IEEE Trans. Microwave Theory Tech.,* vol. 59, pp. 2739-2750, Oct. 2011.

[28] G. Liang, N. Neophytou, D. E. Nikonov, and M. S. Lundstrom, "Performance

projections for ballistic graphene nanoribbon field-effect transistors," *IEEE Trans. Electron Devices,* vol. 54, pp. 677-682, Apr. 2007.

[29] N. Vandecasteele, A. Barreiro, M. Lazzeri, A. Bachtold, and F. Mauri, "Current-voltage characteristics of graphene devices: Interplay between Zener-Klein tunneling and defects," *Phys. Rev. B (Condensed Matter and Materials Physics),* vol. 82, pp. 045416-1-10, Jul. 2010.

[30] AWR Corporation, "Microwave Office," version 9.05r, 2010.

[31] J. Chauhan and J. Guo, "Inelastic phonon scattering in graphene FETs," *IEEE Trans. Electron Devices*, vol. 58, no. 11, pp. 3997-4003, Nov. 2011.

[32] K. N. Parrish and D. Akinwande, "Impact of contact resistance on the transconductance and linearity of graphene transistors," *Appl. Phys. Lett.*, vol. 98, no. 18, pp. 183505-1-3, May. 2011.

[33] S. Rodriguez, A. Smith, S. Vaziri, M. Ostling, M. C. Lemme, and A. Rusu, "Static nonlinearity in graphene field effect transistors," *IEEE Trans. Electron Devices,* vol. 61, no. 8, pp. 3001-3003, Aug. 2014.

[34] Stokholm Sweden: COMSOL Inc., "COMSOL Multiphysics," version 3.4, 2004.

[35] Y. Yoon, G. Fiori, G. Iannaccone, J. Guo, "Performance comparison of graphene nanoribbon FETs with Schottky contacts and doped reservoirs," *IEEE Trans. Electron Devices*, vol. 55, no. 9, pp. 2314-2323, Sep. 2008.

[36] H. Wang, A. Hsu, J. Wu, J. Kong, and T. Palacios, "Graphene-based ambipolar RF mixers," *IEEE Electron Device Lett.*, vol. 31, pp. 906-908, Sep. 2010.

[37] O. Habibpour, J. Vukusic, and J. Stake, "A 30-GHz integrated subharmonic mixer based on a multichannel graphene FET," *IEEE Trans. Microw. Theory Tech.*, vol. 61, pp. 841-847, Feb. 2013.

[38] M. A. Andersson, O. Habibpour, J. Vukusic, and J. Stake, "Resistive graphene FET subharmonic mixers: noise and linearity assessment," *IEEE Trans. Microw. Theory Tech.*, vol. 60, pp. 4035-4042, Dec. 2012.

[39] J. S. Moon, H. -C. Seo, M. Antcliffe, D. Le, C. McGuire, A. Schmitz, L. O. Nyakiti,

D. K. Gaskill, P. M. Campbell, K. -M. Lee, and P. Asbeck, "Graphene FETs for zero-bias linear resistive FET mixers," *IEEE Electron Device Lett.*, vol. 34, pp. 465-467, Mar. 2013.

[40] H. Madan, M. J. Hollander, M. LaBella, R. Cavalero, D. Snyder, J. A. Robinson, and S. Datta, "Record high conversion gain ambipolar graphene mixer at 10 GHz using scaled gate oxide," in *IEEE International Electron Devices Meeting (IEDM)*, USA, Dec. 2012, pp. 4.3.1-4.

[41] H. Madan, M. J. Hollander, J. A. Robinson, and S. Datta, "Analysis and benchmarking of graphene based RF low noise amplifiers," in *Device Research Conference*, USA, Jun. 2013, pp. 41-42.

[42] K. A. Jenkins, D. B. Farmer, S. -J. Han, C. Dimitrakopoulos, S. Oida, and A. Valdes-Garcia, "Linearity of graphene field-effect transistors," *Appl. Phys. Lett.*, vol. 103, no. 17, pp. 173115-1-4, Oct. 2013.

[43] D. O. Pederson and K. Mayaram, *Analog Intergrated Circuit for Communication: Principles, Simulation and Design.* Springer, 2008.

[44] A. Javey, J. Guo, D. B. Farmer, Q. Wang, D. Wang, R. G. Gordon, M. Lundstrom, and H. Dai, "Carbon nanotube field-effect transistors with integrated ohmic contacts and high-k gate dielectrics," *Nano Letters,* vol. 4, pp. 447-450, Mar. 2004.

[45] Z. Chen, D. Farmer, S. Xu, R. Gordon, P. Avouris, and J. Appenzeller, "Externally assembled gate-all-around carbon nanotube field-effect transistor," *IEEE Electron Device Lett.,* vol. 29, pp. 183-185, Feb. 2008.

[46] A. Javey, R. Tu, D. B. Farmer, J. Guo, R. G. Gordon, and H. Dai, "High performance n-type carbon nanotube field-effect transistors with chemically doped contacts," *Nano Letters,* vol. 5, pp. 345-348, Jan. 2005.

[47] I. Amlani, K. F. Lee, J. Deng, and H.-S. Philip Wong, "Measuring frequency response of a single-walled carbon nanotube common-source amplifier," *IEEE Trans. Nanotechnol.,* vol. 8, no. 2, pp. 226-233, Mar. 2009.

[48] R. Kumashiro, Y. Wang, N. Komatsu, and K. Tanigaki, "FET characteristics of

chemically-modified CNT," in *2009 MRS Spring Meeting: MRS Symposium B on Concepts in Molecular and Organic Electronics,* Apr. 2009, pp. 35-39.

[49] K. Ryu, A. Badmaev, C. Wang, A. Lin, N. Patil, L. Gomez, A. Kumar, S. Mitra, H.-S. Philip Wong, and C. Zhou, "CMOS-analogous wafer-scale nanotube-on-insulator approach for submicrometer devices and integrated circuits using aligned nanotubes," *Nano Letters,* vol. 9, pp. 189-197, Jan. 2009.

[50] J. Chen, C. Klinke, A. Afzali, and P. Avouris, "Self-aligned carbon nanotube transistors with charge transfer doping," *Appl. Phys. Lett.,* vol. 86, no. 12, pp. 123108-1-3, Mar. 2005.

[51] A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. Dai, "Ballistic carbon nanotube field-effect transistors," *Nature,* vol. 424, pp. 654-657, Aug. 2003.

[52] S. Moon, S. Lee, W. Song, J. S. Lee, N. Kim, J. Kim, and N. Park, "Fabrication of n-type nanotube transistors with large-work-function electrodes," *Appl. Phys. Lett.,* vol. 90, no. 9, pp. 92113-1-3, Feb. 2007.

[53] F. N. Ishikawa, H. Chang, K. Ryu, P. Chen, A. Badmaev, L. G. De Arco, G. Shen, and C. Zhou, "Transparent electronics based on transfer printed aligned carbon nanotubes on rigid and flexible substrates," *ACS Nano,* vol. 3, pp. 73-79, Jan. 2009.

[54] L. Ding, S. Wang, Z. Zhang, Q. Zeng, Z. Wang, T. Pei, L. Yang, X. Liang, J. Shen, Q. Chen, R. Cui, Y. Li, and L. Peng, "Y-contacted high-performance n-type single-walled carbon nanotube field-effect transistors: Scaling and comparison with Sc-contacted devices," *Nano Letters,* vol. 9, pp. 4209-4214, Dec. 2009.

[55] Z. Zhang, S. Wang, Z. Wang, L. Ding, T. Pei, Z. Hu, X. Liang, Q. Chen, Y. Li, and L. Peng, "Almost perfectly symmetric SWCNT-based CMOS devices and scaling," *ACS Nano,* vol. 3, pp. 3781-3787, Nov. 2009.

[56] A. D. Franklin, G. Tulevski, J. B. Hannon, and Z. Chen, "Can carbon nanotube transistors be scaled without performance degradation?" in *2009 IEEE International Electron Devices Meeting (IEDM),* Dec. 2009, pp. 1-4.

[57] J. Guo, A. Javey, H. Dai, S. Datta, and M. Lundstrom, "Predicted performance

advantages of carbon nanotube transistors with doped nanotubes as source/drain," *arXiv:cond-mat/0309039*, Sep. 2003.

[58] A. Javey and J. Kong, Eds., *Carbon Nanotube Electronics.* Springer US, 2009.

[59] M. A. Wahab and K. Alam, "Performance comparison of zero-Schottky-barrier and doped contacts carbon nanotube transistors with strain applied," *Nano-Micro Lett,* vol. 2, pp. 126-133, Jul. 2010.

[60] D. L. Pulfrey and L. Chen, "Examination of the high-frequency capability of carbon nanotube FETs," *Solid-State Electronics,* vol. 52, pp. 1324-1328, Sep. 2008.

[61] J. Guo, S. Hasan, A. Javey, G. Bosman, and M. Lundstrom, "Assessment of high-frequency performance potential of carbon nanotube transistors," *IEEE Trans. Nanotechnol.,* vol. 4, pp. 715-21, Nov. 2005.

[62] N. Paydavosi, K. D. Holland, M. M. Zargham, and M. Vaidyanathan, "Understanding the frequency- and time-dependent behavior of ballistic carbon-nanotube transistors," *IEEE Trans. Nanotechnol.,* vol. 8, pp. 234-244, Mar. 2009.

[63] B. Razavi, *RF Microelectronics.* Prentice Hall, New York, 1998.

[64] K. Holland, N. Paydavosi, and M. Vaidyanathan, "Self-consistent simulation of array-based CNFETs: Impact of tube pitch on RF performance," in *2010 14th International Workshop on Computational Electronics (IWCE),* Oct. 2010, pp. 1-4.

[65] M. Vaidyanathan, M. Iwamoto, L. E. Larson, P. S. Gudem, and P. M. Asbeck, "A theory of high-frequency distortion in bipolar transistors," *IEEE Trans. Microwave Theory Tech.,* vol. 51, pp. 448-61, Feb. 2003.

[66] S. Datta, *Quantum Transport: Atom to Transistor.* Cambridge University Press, 2005.

[67] Y. Gotoh, K. Mukai, Y. Kawamura, H. Tsuji, and J. Ishikawa, "Work function of low index crystal facet of tungsten evaluated by the Seppen-Katamuki analysis," *J. Vac. Sci. Technol. B (Microelectronics and Nanometer Structures),* vol. 25, pp. 508-512, Mar. 2007.

[68] S. J. Tans, R. M. Verschueren, and C. Dekker, "Room temperature transistor based

on a single carbon nanotube," *Nature,* vol. 393, pp. 49-52, May 1998.

[69] J. Itoh, K. Tsuburaya, S. Kanemaru, T. Watanabe, and S. Itoh, "Fabrication and characterization of comb-shaped lateral field-emitter arrays," *Jpn. J. Appl. Phys., Part 1: Regular Papers and Short Notes and Review Papers,* vol. 32, pp. 1221-1226, Mar. 1993.

[70] T. Isogai, H. Tanaka, T. Goto, A. Teramoto, S. Sugawa, and T. Ohmi, "Formation and property of yttrium and yttrium sillicide films as low Schottky barrier material for n-type silicon," *Jpn. J. Appl. Phys.,* vol. 47, pp. 3138-3141, Apr. 2008.

[71] C. Kocabas, J. K. Seong, T. Ozel, M. Shim, and J. A. Rogers, "Improved synthesis of aligned arrays of single-walled carbon nanotubes and their implementation in thin film type transistors," *J. Phys. Chem. C,* vol. 111, pp. 17879-17886, Dec. 2007.

[72] G. Liang, N. Neophytou, M. S. Lundstrom, and D. E. Nikonov, "Ballistic graphene nanoribbon metal-oxide-semiconductor field-effect transistors: A full real-space quantum transport simulation," *J. Appl. Phys.,* vol. 102, pp. 054307-1-7, Sep. 2007.

[73] G. Fiori and G. Iannaccone, "Simulation of graphene nanoribbon field-effect transistors," *IEEE Electron Device Lett.,* vol. 28, pp. 760-762, Aug. 2007.

[74] T. Low, S. Hong, J. Appenzeller, S. Datta, and M. S. Lundstrom, "Conductance asymmetry of graphene p-n junction," *IEEE Trans. Electron Devices,* vol. 56, pp. 1292-1299, Jun. 2009.

[75] K. Alam, "Transport and performance of a zero-Schottky barrier and doped contacts graphene nanoribbon transistors," *Semicond. Sci. Technol.,* vol. 24, pp. 015007-1-8, Jan. 2009.

[76] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films," *Science,* vol. 306, pp. 666-669, Oct. 2004.

[77] Y. Wu, Y. Lin, A. A. Bol, K. A. Jenkins, F. Xia, D. B. Farmer, Y. Zhu, and P. Avouris, "High-frequency, scaled graphene transistors on diamond-like carbon," *Nature,* vol. 472, pp. 74-78, Apr. 2011.

[78] L. Liao, Y. Lin, M. Bao, R. Cheng, J. Bai, Y. Liu, Y. Qu, K. L. Wang, Y. Huang, and X. Duan, "High-speed graphene transistors with a self-aligned nanowire gate," *Nature,* vol. 467, pp. 305-308, Sep. 2010.

[79] Y. -M. Lin, C. Dimitrakopoulos, K. A. Jenkins, D. B. Farmer, H. -Y. Chiu, A. Grill, and P. Avouris, "100-GHz transistors from wafer-scale epitaxial graphene," *Science,* vol. 327, pp. 662, Feb. 2010.

[80] M. C. Lemme, T. J. Echtermeyer, M. Baus, and H. Kurz, "A graphene field-effect device," *IEEE Electron Device Lett.,* vol. 28, pp. 282-284, Apr. 2007.

[81] Z. Chen, Yu-Ming Lin, M. J. Rooks, and P. Avouris, "Graphene nano-ribbon electronics," *Physica E: Low-Dimensional Systems and Nanostructures,* vol. 40, pp. 228-232, Dec. 2007.

[82] T. G. Pedersen, C. Flindt, J. Pedersen, N. A. Mortensen, A. -P. Jauho, and K. Pedersen, "Graphene antidot lattices: designed defects and spin qubits," *Phys. Rev. Lett.,* vol. 100, pp. 136804-1-4, Apr. 2008.

[83] T. Ohta, A. Bostwick, T. Seyller, K. Horn, and E. Rotenberg, "Controlling the electronic structure of bilayer graphene," *Science,* vol. 313, pp. 951-954, Aug. 2006.

[84] Y. Wu, K. A. Jenkins, A. Valdes-Garcia, D. B. Farmer, Y. Zhu, A. A. Bol, C. Dimitrakopoulos, W. Zhu, F. Xia, P. Avouris, and Y. Lin, "State-of-the-art graphene high-frequency electronics," *Nano Lett.*, vol. 12, pp. 3062-3067, May 2012.

[85] Y. -M. Lin, A. Valdes-Garcia, S. Han, D. B. Farmer, I. Meric, Y. Sun, Y. Wu, C. Dimitrakopoulos, A. Grill, P. Avouris, and K. A. Jenkins, "Wafer-scale graphene integrated circuit," *Science,* vol. 332, pp. 1294-1297, Jun. 2011.

[86] J. Kang, D. Sarkar, Y. Khatami, and K. Banerjee, “Proposal for all-graphene monolithic logic circuits,” *Appl. Phys. Lett.*, vol. 103, no. 8, pp. 083113-1-5, Aug. 2013.

[87] S. Kim, J. Nah, I. Jo, D. Shahrjerdi, L. Colombo, Z. Yao, E. Tutuc, and S. K. Banerjee, "Realization of a high mobility dual-gated graphene field-effect transistor with Al_2O_3 dielectric," *Appl. Phys. Lett.*, vol. 94, no. 6, pp. 062107-1-3, Feb. 2009

[88] D. B. Farmer, Y. Lin, and P. Avouris, "Graphene field-effect transistors with self-aligned gates," *Appl. Phys. Lett.*, vol. 97, no. 1, pp. 013103-1-3, Jul. 2010

[89] D. Im, I. Nam, and K. Lee, "A CMOS active feedback balun-LNA with high IIP2 for wideband digital TV receivers," *IEEE Trans. Microw. Theory Tech*, vol. 58, no. 12, pp. 3566-3579, Dec. 2010.

[90] F. Xia, V. Perebeinos, Y. Lin, Y. Wu, and P. Avouris, "The origins and limits of metal-graphene junction resistance," *Nat. Nano.,* vol. 6, no. 6, pp. 179-184, Mar. 2011.

[91] J. S. Moon, M. Antcliffe, H. C. Seo, D. Curtis, S. Lin, A. Schmitz, I. Milosavljevic, A. A. Kiselev, R. S. Ross, D. K. Gaskill, P. M. Campbell, R. C. Fitch, K. -M. Lee, and P. Asbeck, "Ultra-low resistance ohmic contacts in graphene field effect transistors," *Appl. Phys. Lett.*, vo. 100, no. 20, pp. 203512-1-3, May 2012.

[92] F. M. Kin, C. Lee, J. Hone, J. Shan, and T. F. Heinz, "Atomically thin MoS_2: A new direct-gap semiconductor," *Phys. Rev. Lett.,* vol. 105, pp. 136805-1-4, Sep. 2010.

[93] A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C. Chim, G. Galli, and F. Wang, "Emerging photoluminescence in monolayer MoS_2," *Nano Letters,* vol. 10, pp. 1271-1275, Apr. 2010.

[94] J. Lee, H. Chang, K. N. Parrish, H. Li, R. S. Ruoff, and D. Akinwande, "State-of-the-art flexible 2D nanoelectronics based on graphene and MoS_2," in *71St Device Research Conference (DRC),* Jun. 2013, pp. 43-44.

[95] B. Radisavljevic, M. B. Whitwick, and A. Kis, "Small-signal amplifier based on single-layer MoS_2," *Appl. Phys. Lett.,* vol. 101, pp. 043103-1-4, Jul. 2012.

[96] Y. Yoon, K. Ganapathi, and S. Salahuddin, "How good can monolayer MoS_2 transistors be?" *Nano Letters,* vol. 11, pp. 3768-3773, Sep. 2011.

[97] B. Radisavljevic, M. B. Whitwick, and A. Kis, "Small-signal amplifier based on single-layer MoS_2," *Appl. Phys. Lett.,* vol. 101, pp. 043103-1-4, Jul. 2012.

[98] H. Wang, L. Yu, Y. Lee, Y. Shi, A. Hsu, M. L. Chin, L. Li, M. Dubey, J. Kong, and T. Palacios, "Integrated circuits based on bilayer MoS_2 transistors," *Nano Letters,*

vol. 12, pp. 4674-4680, Sep. 2012.

[99] H. Wang, L. Yu, Y. Lee, W. Fang, A. Hsu, P. Herring, M. Chin, M. Dubey, L. Li, J. Kong, and T. Palacios, "Large-scale 2D electronics based on single-layer MoS_2 grown by chemical vapor deposition," in *2012 IEEE International Electron Devices Meeting (IEDM),* Dec. 2012, pp. 4.6.1-4.6.4.

[100] S. Das and J. Appenzeller, "On the importance of bandgap formation in graphene for analog device applications," *IEEE Trans. Nanotechnol.,* vol. 10, pp. 1093-1098, Sep. 2011.

[101] L. Liu, S. B. Kumar, Y. Ouyang, and J. Guo, "Performance limits of monolayer transition metal dichalcogenide transistors," *IEEE Trans. Electron Devices,* vol. 58, pp. 3042-3047, Sep. 2011.

[102] M. S. Lundstrom, *Fundamentals of Carrier Transport.* Cambridge University Press, New York, 2009.