Highly Accurate SPICE-Compatible Modeling for Single- and Double-Gate GNRFETs with Studies on Technology Scaling

Morteza Gholipour[1,2], Ying-Yu Chen[1], Amit Sangai[1], and Deming Chen[1]
[1]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL
[2]Advanced VLSI Lab., School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
{morteza, chen150, sangai2, dchen}@illinois.edu

***Abstract*—In this paper, we present a highly accurate closed-form compact model for Schottky-Barrier-type Graphene Nano-Ribbon Field-Effect Transistors (SB-GNRFETs). This is a physics-based analytical model for the current–voltage (I–V) characteristics of SB-GNRFETs. We carry out accurate approximations of Schottky barrier tunneling, channel charge and current, which provide improved accuracy while maintaining compactness. This SPICE-compatible compact model surpasses the existing model [15] in accuracy, and enables efficient circuit-level simulations of futuristic GNRFET-based circuits. The proposed model considers various design parameters and process variation effects, including graphene-specific edge roughness, which allows complete and thorough exploration and evaluation of SB-GNRFET circuits. We are able to model both single- and double-gate SB-GNRFETs, so we can evaluate and compare these two types of SB-GNRFET. We also compare circuit-level performance of SB-GNRFETs with multi-gate (MG) Si-CMOS for a scalability study in future generation technology. Our circuit simulations indicate that SB-GNRFET has an energy-delay product (EDP) advantage over Si-CMOS; the EDP of the ideal SB-GNRFET (assuming no process variation) is ~1.3% of that of Si-CMOS, while the EDP of the non-ideal case with process variation is 136% of that of Si-CMOS. Finally, we study technology scaling with SB-GNRFET and MG Si-CMOS. We show that the EDP of ideal (non-ideal) SB-GNRFET is ~0.88% (54%) EDP of that of Si-CMOS as the technology nodes scales down to 7 nm.**

I. Introduction

Graphene has recently received a lot of attention as a material for nanoelectronic devices due to its outstanding physical and electrical properties and its potential to replace Si for building transistors [1]-[3]. The thin, planar, and robust lattice makes graphene potentially compatible with the existing Si-CMOS manufacturing technology [4] and suitable for making flexible electronics [5]. Meanwhile, successfully fabricated devices have been demonstrated [4]-[8], where the fabricated graphene nano-ribbons (GNRs) can have widths below 10 nm and of high quality with fairly smooth edges. In particular, a Si compatible, transfer-free, and *in situ* GNRFET fabrication method is presented in [4], demonstrating compatibility and integrability of GNRFETs with the existing Si technology available in the industry. One advantage in the SB-GNRFET is that it requires no doping in the terminals or the channel. Therefore, it reduces the technical difficulty in fabrication and eliminates doping variation. As a result, most fabricated GNRFETs reported in literature are SB-type [4]-[8].

Modeling and computer simulation are very useful in providing physical insights of GNRFETs and evaluating the performance of futuristic graphene-based circuits. Numerical simulations based on non-equilibrium Green's function (NEGF) formalism have been implemented in the 3-D device simulator NanoTCAD ViDES in [9]-[11]. Non-closed-form analytical models that describe SB-GNRFETs are presented in [12] and [13]. In terms of high-level simulations, a circuit simulation framework of SB-GNRFETs based on lookup tables is presented in [14]. Due to the complicated tunneling effects occurring at the Schottky barriers, no physics-based closed-form model of SB-GNRFETs has been developed until the work [15]. In [15], a compact model of SB-GNRFETs based on a cut-off approximation of the Schottky barrier tunneling probability is presented. However, this is an oversimplified approximation, and the model is inaccurate in the region where the Schottky barrier tunneling effect is prominent. Therefore, the model in [15] results in an overestimation of the OFF current I_{off}, giving a pessimistic view of the I_{on}/I_{off} ratio. The work [15] mainly focused on single-gate SB-GNRFET designs, which had less gate control on the channel as compared to double-gate designs, and thus aggravated the problem.

In this work, we develop a physics-based analytical model for the current–voltage (I–V) characteristics of SB-GNRFETs that is based on elaborative approximations of Schottky barrier tunneling, which provides improved accuracy over the model in [15] while maintaining compactness. In addition, we focus on both single-gate (SG) and double-gate (DG) designs in this work as opposed to the work [15], which emphasized on SG designs and demonstrated that SB-GNRFETs did not have a good I_{on}/I_{off} ratio. DG transistors have a better gate control over the channel and hence a better I_{on}/I_{off} ratio. With the proposed model, we provide more accurate and realistic simulations of SB-GNRFET circuits than the work [15].

For a fair comparison and for the increasing trend in the use of multi-gate (MG) transistors, we compare with MG Si-CMOS designs (e.g. FinFETs) [16] in our circuit simulation experiments. Note that DG graphene-based transistors are fabricated in a planar fashion due to graphene's thin-film structure [17], while MG Si-CMOS transistors are usually of a 3-D FinFET-like structure. Nevertheless, they both demonstrate better gate control ability than SG designs and are likely to be adopted in the upcoming technology nodes.

With the proposed model, we perform a comparative study on SG and DG SB-GNRFETs with MG Si-CMOS on their respective circuit-level delay and power performance. Because GNRFET is regarded as a next-generation device, we are interested in its scalability in future technology nodes. Therefore, we simulate benchmark circuits on the 16 nm, 14 nm, 10 nm, and 7 nm technology nodes to provide insights on scalability. We show that SB-GNRFET circuits have a consistently decreasing trend in delay, power, and EDP with respect to the transistor size, indicating that SB-GNRFET is a promising device in future technology nodes.

To summarize, the main contributions of this work are:

- Proposing an effective and detailed closed-form approximation of the Schottky-barrier tunneling effect.
- Developing a highly accurate compact SB-GNRFET model, supporting both SG and DG transistor designs.
- Comparing circuit-level performance among MG Si-CMOS, single-gate SB-GNRFET, and double-gate SB-GNRFET.
- Providing insights on technology scaling with the above technology nodes.

The rest of the paper is organized as follows: Section II covers more background knowledge on graphene and GNRs. Section III discusses the modeling of tunneling in SB-GNRFETs and presents our SB-GNRFET compact model. Section IV and V present the experimental results, including model validation and circuit simulations, and Section VI concludes the paper.

978-3-9815370-2-4/DATE14/©2014 EDAA

II. Graphene Energy Dispersion

Graphene is a single atomic layer of graphite with two-dimensional honeycomb crystal lattice. It is a zero-band-gap material, which makes it metallic and unable to be turned on or off [18]. Energy gap can however be induced by means of lateral confinement [19]. In order to open the band gap and make graphene into a good semiconductor, it is patterned into 1-D GNRs with widths below 10 nm [18]. The band gap of a GNR is mainly inversely proportional to its width [20]. The width of a GNR (denoted by W_{CH}) is commonly defined via the number of dimer lines N as illustrated in Fig. 1 (a), as $W_{CH} = \sqrt{3}d_{cc}\,(N+1)/2$, where d_{cc} = 0.142 nm refers to the carbon–carbon bond distance [21]. As the width of the GNR increases, the band structure of GNRs gradually returns to that of a 2-D graphene sheet.

Based on the edge geometry, GNRs are categorized into two types: armchair-GNRs (AGNR) and zigzag-GNRs (ZGNR) [22]. In this work, we focus on AGNRs due to its semiconducting property. The energy dispersion relation of an AGNR for subband $\alpha = 1,2,\ldots,N$ is given in [23]-[24] as

$$E_\alpha(k) = \pm\tau\sqrt{1 + 4\,A_\alpha \cos\frac{\sqrt{3}a_l k}{2} + 4{A_\alpha}^2} \tag{1}$$

where k is the wavevector, $A_\alpha = \cos(\pi\alpha/(N+1))$, $a_l = \sqrt{3}d_{cc}$, and $\tau = 2.7$ eV is the nearest neighbor overlap energy. The latter parameter is different for the carbon atoms at the edge of the ribbon. This can be accounted by the edge-corrected energy dispersion $E_\alpha^c(k) = E_\alpha(k) + E_\alpha^\delta(k)$, in which the correction energy is obtained using the approach given in [20] as

$$E_\alpha^\delta(k) = s_\alpha \frac{4\upsilon\tau}{N+1}\sin^2\left(\frac{\alpha\pi}{N+1}\right)\cos(ka_l) \tag{2}$$

where $\upsilon = 0.12$ eV, and

$$s_\alpha = \begin{cases} 1 & A_\alpha \geq -1/2 \\ -1 & \text{otherwise} \end{cases}$$

The density of states (DOS) can be obtained from the effective-mass (EM) approximation as

$$D_\alpha(E) = (2/\pi\hbar)\sqrt{M_\alpha/2E} \tag{3}$$

where $\hbar$ is the reduced Planck's constant, $E = E_\alpha^c(k) - \varepsilon_\alpha$ is the energy with respect to the band edge energy $\varepsilon_\alpha = E_\alpha^c(0)$, and M_α is the effective mass given by

$$M_\alpha = -\frac{2\hbar^2\varepsilon_\alpha}{3{a_l}^2\tau^2 A_\alpha} \tag{4}$$

III. SB-GNRFET Modeling

An SB-GNRFET consists of a GNR-based channel and metal electrodes: gate, drain, and source. An example of SB-GNRFET is given in Fig. 1 (b). The interface between the metal drain/source and the GNR channel results in a Schottky barrier at the graphene-metal junction. SB-GNRFETs have an ambipolar I-V curve with minimum current at V_{GS} = ½ V_{DS} [12]-[13]. Multiple GNRs can be connected in parallel to increase driving strength, as in Fig. 2 (a).

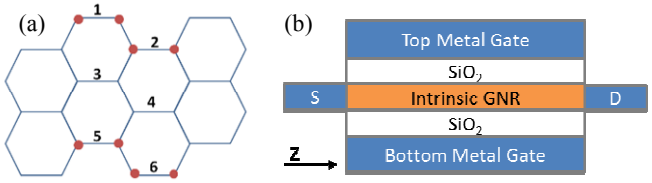

Fig. 1. (a) Lattice structure of an AGNR with N=6. N is the number of dimer lines in the armchair orientation, (b) Cross section of an SB-GNRFET device.

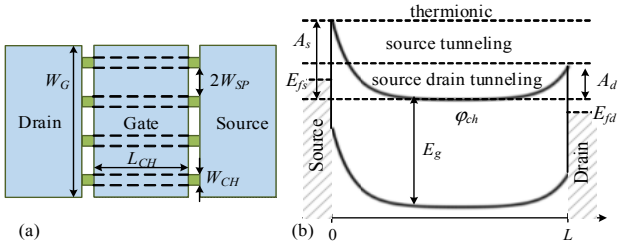

Fig. 2. (a) The structure of a four-ribbon SB-GNRFET. A common drain and a common source are shared by the ribbons. (b) Spatial band diagram along the transport direction of SB-GNRFET.

The GNRFET of interest has the following design parameters: L_{CH} is channel length, W_{CH} is the ribbon width, W_G is the gate width, $2W_{sp}$ is the ribbon spacing, and T_{ox} is the oxide thickness.

A. Schottky Barrier and Tunneling

Schottky barriers are introduced on the interface of metal and graphene. With Schottky barriers present, the charge transport in the device is dominated by Schottky barrier tunneling. The Schottky barrier width is modulated by the gate voltage, changing the tunneling probability for carriers. The band diagram of SB-GNRFET has three distinctive regions: two injecting regions at the ends of the GNR and a central region where ballistic transport occurs (Fig. 2 (b)). In a sufficiently-long-channel device, the central region is of flat-band type with the electrostatic potential of φ_{ch}.

The Schottky barrier profile near the metal/GNR interface, which can be analytically solved from the 1-D Laplace equation, takes an exponential decaying form [12] along the channel direction z. The Schottky barrier profile near the source interface is given by

$$E_{SB}(z) = A_s e^{-z/\lambda} \tag{5}$$

where $\lambda = 2t_{ox}/\pi$ is the scale length, and $A_s = q\varphi_{ch}$ for the lowest subband at the source. The valence band has the same profile as the conduction band but is downshifted by an amount equal to the GNR energy gap $E_g = 2\varepsilon_\alpha$. In the case of $|\varphi_{ch}| > E_g$, the spatial band diagram curvature becomes high enough to trigger band-to-band tunneling (BTBT). In this case, a carrier with energy $0 < E < A_s - 2\varepsilon_\alpha$ experiences a Schottky barrier of a height $A_s = E + 2\varepsilon_\alpha$.

The tunneling phenomenon is characterized by the transmission coefficient $T(E)$ of a carrier. In the case of thermionic conduction, in which the carrier has higher energy than the Schottky barrier, $T(E)$ is equal to unity. For the tunneling conduction, the transmission through a single Schottky barrier is computed based on the Wentzel–Kramers–Brillouin (WKB) approximation [25]

$$T(E) = \exp\left\{-2\int_{z_1}^{z_2}\mathrm{Im}[k_z(E)]dz\right\} \tag{6}$$

where $z_1 = 0$ and $z_2 = -\lambda\ln(E/A_s)$ are the classical turning points, and $\mathrm{Im}[k_z(E)]$ is the imaginary part of the wavevector. The momentum $k_z(E)$ is related to the energy through the GNR E–k dispersion relationship of (1), which can be obtained by expanding $\cos(x) \backsimeq (1 - x^2/2)$ as

$$k_z(E) \backsimeq \sqrt{M_\alpha/\hbar^2\varepsilon_\alpha}\sqrt{{E_\alpha}^2 - {\varepsilon_\alpha}^2} \tag{7}$$

Inserting barrier profile of $E_{SB}(z)$ into this equation results in

$$\mathrm{Im}[k_z(E)] \backsimeq \sqrt{M_\alpha/\hbar^2\varepsilon_\alpha}\sqrt{{\varepsilon_\alpha}^2 - (E + \varepsilon_\alpha - A_s e^{-z/\lambda})} \tag{8}$$

Integrating for $E < A_s$ leads to the transmission coefficient of

$$T(E) = \exp\left\{-2\lambda\sqrt{M_\alpha/\hbar^2\varepsilon_\alpha}\left[(E+\varepsilon_\alpha)\left(\frac{\pi}{2} - \mathrm{atan}\frac{\varphi_\alpha}{\gamma_1}\right) + \gamma_1 + \gamma_2\left(\mathrm{atan}\left(\frac{\gamma_1\gamma_2}{A_s(E+\varepsilon_\alpha) - E(E+2\varepsilon_\alpha)}\right) - \theta_0\right)\right]\right\} \tag{9}$$

where

$$\gamma_1 = \sqrt{{\varepsilon_\alpha}^2 - {\varphi_\alpha}^2}, \qquad \gamma_2 = \sqrt{(\varphi_\alpha + A_s)^2 - {\varepsilon_\alpha}^2}$$
$$\varphi_\alpha = \varepsilon_\alpha + E - A_s, \qquad \theta_0 = \begin{cases} \pi & E(2\varepsilon_\alpha + E) < A_s(\varepsilon_\alpha + E) \\ 0 & otherwise \end{cases} \tag{10}$$

B. Full SB-GNRFET Model

The equivalent circuit of the full GNRFET as in Fig. 2 (a), is shown in Fig. 3 (a). It consists of multiple parallel GNRs and parasitic capacitors C_{gd} and C_{gs}. Each transistor symbol marked in red represents a single GNR and is modeled by the circuit in Fig. 3 (b). In a single GNR, I_{ds} models the current flowing through the channel, the capacitors $C_{ch,d}$, $C_{ch,s}$, $C_{g,ch}$, and $C_{sub,ch}$ model the parasitics, and the voltage-controlled voltage source V_{ch} represents the channel voltage that corresponds to the channel potential φ_{ch}, expressed as $V_{ch} = q\varphi_{ch}$. The capacitors C_{gd} and C_{gs} are modeled based on FastCap [26], which are functions of W_G and T_{ox}, as

$$C_{gd} = C_{gs} = 1.26 \times 10^{-10} W_G (0.8 - 0.2T_{ox} + 0.015T_{ox}^2) \tag{11}$$

Intrinsic capacitors $C_{ch,d} = \partial Q_{ch}/\partial V_D$ and $C_{ch,s} = \partial Q_{ch}/\partial V_S$ are implemented in SPICE as voltage-controlled capacitors by defining the charge equations. The total channel charge Q_{ch} is derived from the electron and hole density of each subband coming from the drain and the source. Depending on the magnitude of the applied bias, multiple reflections can arise between the series combination of source and drain Schottky barriers. The total charge of carriers subband α can be expressed as

$$Q_\alpha^i(\varphi_{ch}) = q \int D_\alpha(E) \cdot G_Q^i(E) dE \tag{12}$$

where $G_Q^i(E)$ is defined as

$$G_Q^i(E) = T_{TS}^i(E) \cdot f\left(E - E_{\alpha,s}^i\right) + T_{TD}^i(E) \cdot f\left(E - E_{\alpha,d}^i\right) \tag{13}$$

where i can be either e or h, representing total electron or hole charge in the subband α, q is the electron charge, and $f(\cdot)$ is the Fermi–Dirac distribution function given by

$$f(E - E_F) = \left[1 + \exp\left(\frac{E - E_F}{k_b T}\right)\right]^{-1}$$
$$E_{\alpha,j}^e = -\left(\varepsilon_\alpha - q\varphi_{ch} + V_j\right) \qquad j = s, d$$
$$E_{\alpha,j}^h = -\left(-\varepsilon_\alpha + q\varphi_{ch} - V_j\right) \qquad j = s, d \tag{14}$$

where k_b is the Boltzmann constant, T is the temperature in Kelvin, and j=s, d denotes the source and drain terminals, respectively. The tunneling coefficients within this formalism play a critical role as

$$T_{TS}(E) = \frac{T_s(2 - T_d)}{T_s + T_d - T_S T_d}, \quad T_{TD}(E) = \frac{T_d(2 - T_s)}{T_s + T_d - T_S T_d} \tag{15}$$

where T_s (T_d) is the transmission coefficient of a carrier going through the Schottky barrier on the source (drain) side, given by (9). Note that T_s (T_d) should be computed for both electrons and holes. The total mobile charge $Q_{ch} = \sum_\alpha (Q_\alpha^h - Q_\alpha^e)$ must be equal to the charge Q_{cap} across the gate, source and drain capacitors that couple into the channel

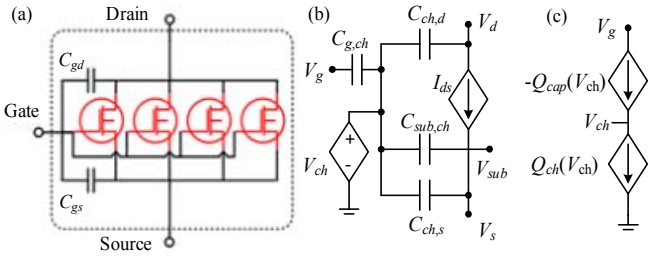

Fig. 3. (a) SPICE model of an SB-GNRFET with 4 parallel GNRs. (b) SPICE model of a single GNR. (c) SPICE setup for solving V_{ch}. V_g, V_d, V_s, and V_{sub} are voltages of each terminal.

$$Q_{cap}(\varphi_{ch}) = -\sum_{i=g,s,d} C_{i,ch}(V_i - V_{FB,i} - q\varphi_{ch}) \tag{16}$$

$$C_{g,ch} = \frac{5.55 \times 10^{-11} \epsilon_r L_{CH}}{\left(1 + \frac{1.5T_{ox}}{W_G}\right) \ln\left(\frac{5.98W_{CH}}{0.8T_{ox}}\right)}$$
$$C_{s,ch} = C_{d,ch} = 0.1C_{g,ch} \tag{17}$$

where $V_{FB,i}$ is the flatband voltage and ϵ_r is the relative permittivity of the material. Equation (17) is also modeled empirically from data extracted from FastCap. Both Q_{ch} and Q_{cap} are functions of φ_{ch} and have to be equal in magnitude due to charge conservation. As a result, equating Q_{ch} and Q_{cap} yields a solution of φ_{ch}, which can be obtained using the equation solver circuit of Fig. 3(c) [27] in SPICE simulations. In the solver, two voltage-controlled current sources are connected in series, forcing V_{ch} to take a value such that the currents from the two sources are equal.

As $T(E)$ in (9) takes a complicated form, it often results in nonclosed-form solutions when integrating with other quantities, making the model non-compact. We made an analytical approximation which simplifies the integration in (12) and makes it closed-form. Following is the explanation of the approximation for the total electron charge. The same method can be employed to compute the total hole charge.

The term G_Q^e of electrons for different cases is shown in Fig. 4. For all the cases, G_Q^e can be approximated by using two lines defined by four values E_{c1}, $E_{c2,Q}$, $G_{Q,0}$ and $G_{Q,1}$. However, it introduces some error for high gate voltages (see Fig. 4 (c) and (d)). These approximations are shown as dashed lines in Fig. 4. The term E_{c1} is determined by the Fermi distribution of source or drain terminals, whichever has a lower voltage, and the corresponding transmission coefficient which is in its linear region. Therefore, G_Q^e would be the product of a linear function mE and the Fermi distribution as

$$G_Q^e \simeq mE \cdot \left[1 + \exp\left(\frac{E - E_F}{k_b T}\right)\right]^{-1}$$
$$E_f = -(\varepsilon_\alpha - q\varphi_{ch} + min\,(V_s, V_d)) \tag{18}$$

By differentiating, equating to zero, and solving, we obtain E_{c1} as

$$E_{c1}(E_f, T) = k_b \cdot T \left[W\left(\exp\left(\frac{E_f}{k_b T} - 1\right) + 1\right)\right] \tag{19}$$

where $W(\cdot)$ is the Lambert W product logarithm function. Fig. 5 shows E_{c1} as a function of E_f for different temperatures. E_{c1} can be approximated as constant, parabolic, or linear functions of E_f in three different regions defined by $x_1 = -0.05$, and $x_2 = 0.145$.

$$E_{c1}(E_f, T) = \begin{cases} k_b T & E_f < -0.05 \\ p_1 {E_F}^2 + p_2 E_F + p_3 & -0.05 < E_f < 0.145 \\ p_4 E_F + p_5 & E_f > 0.145 \end{cases} \tag{20}$$

where $p_i, i = 1,2,\ldots,5$, are temperature-dependent coefficient as $p_i = \eta_1 \cdot T + \eta_2$. Values of η_1 and η_2 obtained by curve fitting are given in Table I. The term $G_{Q,1}$ is computed from (13) and the previously obtained E_{c1}. The term $E_{c2,Q}$ is approximated by using $f\left(E_{c2,Q} - E_f\right) \simeq G_{Q,1}/10$. The integration of (12) can be therefore analytically computed as

$$Q_\alpha^e = q \int_0^{E_{c1}} D_\alpha(E) \cdot (a_1 E + b_1) dE + q \int_{E_{c1}}^{E_{c2,Q}} D_\alpha(E) \cdot (a_2 E + b_2) dE$$
$$= \frac{4}{3} \frac{q\sqrt{2M_\alpha}}{\pi\hbar} \left[\sqrt{E_{c1}} \cdot G_{Q,0} + \frac{G_{Q,1} \cdot E_{c2,Q}}{\sqrt{E_{c1}} + \sqrt{E_{c2}}}\right] \tag{21}$$

The same method can be used to compute the hole charge by using $E_f = -(-\varepsilon_\alpha + q\varphi_{ch} - \max(V_s, V_d))$.

Table I. Values of temperature-dependent coefficients.

	p_1	p_2	p_3	p_4	p_5
η_1	-0.0041	-1.33e-4	1.016e-4	-4.47e-4	6.29e-5
η_2	3.4092	0.2827	0.0035	1.0082	-0.0315

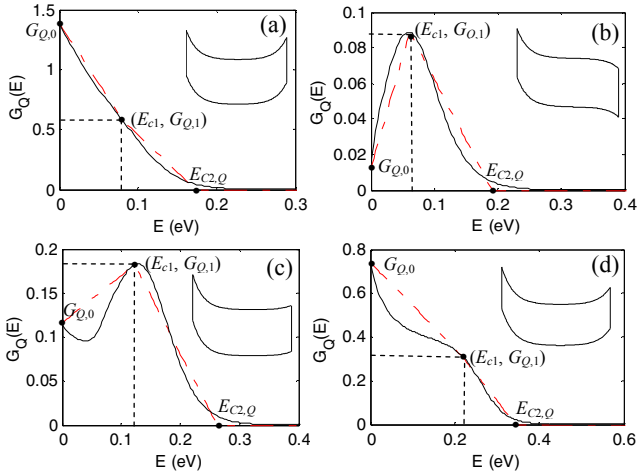

Fig. 4. G_Q^e of electrons for different bias voltages. (a) V_d=0.1 V_g=0.8, (b) V_d=0.5 V_g=0.4, (c) V_d=0.4 V_g=0.55, (d) V_d=0.3 V_g=0.9. Solid lines are numerical and the dashed lines are closed-form analytical. Inset in each subfigure is the energy band diagram in such condition.

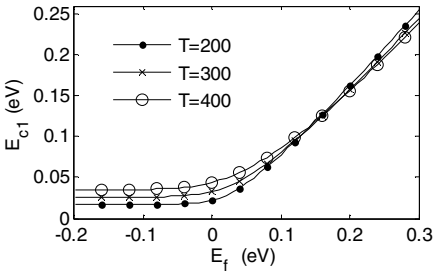

Fig. 5. E_{c1} as a function of E_f for different temperatures T.

Given φ_{ch}, the current through the channel is computed by using the Landauer–Büttiker formalism [28] as

$$I_\alpha^i = \frac{q}{\pi\hbar}\int G_I^i(E)dE \tag{22}$$

$$G_I^i(E) = T_I{}^i\left[f\left(E - E_{\alpha,s}^i\right) - f\left(E - E_{\alpha,d}^i\right)\right] \tag{23}$$

$$T_I = \frac{T_s T_d}{T_s + T_d - T_s T_d} \tag{24}$$

Note that T_I should be computed for both electrons and holes. Here, we are again facing the complexity of integrating G_I^i into closed-form expressions. The term G_I^i can be approximated in the same way as G_Q^i. The term E_{c1} takes the same value obtained for charge approximation given in (19). The term $G_{I,1}$ is computed from (23) for $E = E_{c1}$, and $E_{c2,I}$ is calculated by using $f\left(E_{c_{2,I}} - E_f\right) = G_{I,1}/10$. The current of electrons then can be analytically computed by using the integration of (22) as

$$I_\alpha^e = \frac{q}{\pi\hbar}\left[\int_0^{E_{c1}}(a_1{}'E + b_1{}')dE + \int_{E_{c1}}^{E_{c2,I}}(a_2{}'E + b_2{}')dE\right] = \frac{q}{\pi\hbar}\frac{1}{2}\left[G_{I,0}\cdot E_{c1} + G_{I,1}\cdot E_{c2,I}\right] \tag{25}$$

The same method can be used to compute the current of holes. The total current is given by $I_{DS} = I^e - I^h$.

Existing fabrication technology tends to produce GNRs with imperfect edges, which affects the quantum effects occurring inside GNRs. It is called edge roughness, and is characterized by p_r, the probability that any atom on the edges of a GNR is removed [9]. There are two main effects from edge roughness: 1) Change in effective subbands 2) Disruption in ballistic transport. They are modeled in [15] as

$$I_{rough} = A \cdot I_{DS}(\varepsilon_{\alpha,eff}) \tag{26}$$

where A is the scattering coefficient, and $\varepsilon_{\alpha,eff}$ is the effective subband. We follow a similar scheme to adjust the current with edge roughness based on [15].

IV. **Transistor Model Validation**

In order to validate the proposed model, we compare the I-V curves obtained from the proposed analytical model with results from NanoTCAD ViDES [9]-[11]. We also compare with the model proposed in [15] under the same design parameters and with DG structure in order to demonstrate improvement in accuracy. The I-V curves for design parameter N=12 and $0 < V_g < 0.75$ are plotted in Fig. 6, with T_{ox}=2 nm in (a) and T_{ox}=1.5 nm in (b). Fig. 6 (c) shows the effect of N on the current. It is shown that our model agrees very well with ViDES results and has improved accuracy as compared to the model in [15]. In particular, the proposed model gives a more realistic I_{on}/I_{off} ratio than the model in [15] (e.g. 52% error in I_{on}/I_{off} of Fig. 6 (a) at V_d=0.6), making the subsequent circuit simulations in the next section more representative. Furthermore, our model is much faster than the ViDES simulator, which is based on time-consuming numerical computations. For example, a 10 ns transient simulation of a *nand2* gate lasts about 2 seconds to run, while about 10 minutes is needed to find a single bias point of a SB-GNRFET in ViDES on the same machine.

The minimum current in SB-GNRFETs occurs at V_{GS} = ½ V_{DS}, making the transistor ambipolar in the operating region. This minimum point, which is called *ambipolar conduction point*, occurs at V_{GS} = ½ V_{DS} for a mid-gap Schottky barrier height. Ideally, the minimum current should occur in the OFF state when V_{GS}=0. The minimum current point can be shifted to a different V_{GS} by tuning the gate work function by using different metals [14][29]. In this work, we shift 0.25 V for P-type and -0.25 V for N–type transistors for ideally balanced device strength and performance.

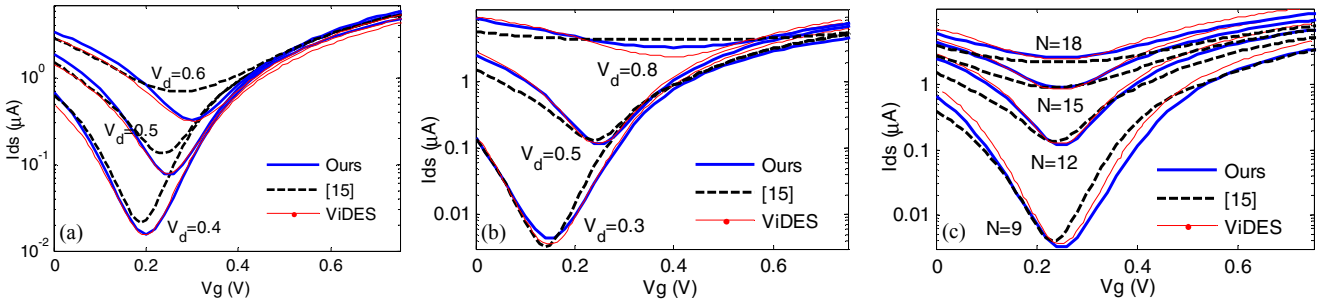

Fig. 6. I-V curves of double-gate (DG) SB-GNRFET, (a) N=12, and T_{ox}=2 nm, (b) N=12, and T_{ox} =1.5 nm, (c) V_d=0.5V, and T_{ox}=1.5 nm.

Here, we define $I_{on} = I(V_{GS} = 0.75, V_{DS} = 0.5)$ and $I_{off} = I(V_{GS} = 0.25, V_{DS} = 0.5)$ by assuming V_{DD}=0.5 and an I-V curve shifting of ±0.25 V. In [20], a periodic effect on band gaps with respect to N is reported, and our model tracks the periodicity very well. For N = 8, 11, 14, and 17, the band gap is very small, resulting in a low I_{on}/I_{off} ratio. For N = 6, 9, 12, 15, and 18, there is a moderate band gap, resulting in a high I_{on}/I_{off} ratio and a high I_{on}. For N = 7, 10, 13, and 16, the band gap is the largest, resulting in the highest I_{on}/I_{off} ratio. However, I_{on} is lower because the large band gaps prevent carriers from occupying the channel. Also, the I_{on}/I_{off} ratio tends to decrease as N increases.

V. Circuit-Level Evaluation

We implemented the proposed model in HSPICE as a library for both SG and DG SB-GNRFET. To evaluate SB-GNRFET circuits, we built digital logic circuits with the implemented library. We also built circuits with the MG Si-CMOS High-Performance (HP) library from Predictive Technology Model (PTM) [30] as a comparison to Si-based technology. We choose the high-performance Si-CMOS library over the low-power one because the SB-GNRFET is known to be a high-performance rather than a low-power device, according to the study in [15]. We implemented benchmark circuits *c17* from ISCAS '85, *b02* from ITC '99, *s27* from ISCAS '89, and carry generator for the third bit of a carry look-ahead adder (*cla*) in the following five technologies: High Performance MG Si-CMOS from PTM, ideal DG SB-GNRFET, non-ideal DG SB-GNRFET, ideal SG SB-GNRFET, and non-ideal SG SB-GNRFET. For the non-ideal cases, we set the edge roughness probability p_r = 0.1 to account for the imperfection from process variation that is unlikely to be avoided in practice. On the other hand, the ideal cases give an optimistic insight on how well GNRFET circuits may perform once fabrication technology becomes more mature. We performed delay and power analysis on these technology nodes, as reported in the following subsections.

A. Supply Voltage Scaling

We first investigate the effects of supply voltage (V_{DD}) scaling on SB-GNRFETs. The delay, power, and energy-delay product (EDP) are reported in Fig. 7. We show that the delay is nearly constant across different supply voltages, but power scales down as V_{DD} decreases. As a result, the EDP also improves as V_{DD} decreases. This indicates that the SB-GNRFET has good potential in terms of low V_{DD} computing. However, it should be noted that lower V_{DD} also lowers the noise margin of the circuit, which should be taken into consideration when choosing the V_{DD}. As a result, we choose V_{DD} =0.5V as the nominal supply voltage in subsequent experiments. We also show that non-ideal SB-GNRFETs with process variation result in a large increase in the delay and EDP. In addition, we show that DG and SG SB-GNRFETs do not have significant difference in terms of circuit-level performance. This is because the I-V curves of SG and DG do not differ by a lot. The I_{on} is ~30% different, while the I_{off} is only ~5% different. Given the higher cost of manufacturing DG designs and the limited performance advantage, it may not be always desirable to prefer DG designs over SG.

B. Cross-Technology Comparison

Simulation results of benchmark circuits are presented in Fig. 8. We report maximum delay, dynamic power, leakage power, total power, and EDP for all the circuits. The trends of delay, power, and EDP are mostly consistent across different circuits. We show that ideal SB-GNRFET, either SG or DG, has lower delay than Si-CMOS (30% or 24% of Si-CMOS). SG SB-GNRFET has slightly lower dynamic power than DG SB-GNRFET due to its lower I_{on}. They have comparable leakage power due to very similar I_{off}. However, since they both have lower I_{on}/I_{off} ratios than Si-CMOS, their leakage power is higher than that of Si-CMOS. In terms of total power dissipation and EDP, SB-GNRFET outperforms Si-CMOS significantly. Ideal DG (SG) SB-GNRFET consumes only 21.9% (19.0%) total power, while non-ideal DG (SG) SB-GNRFET consumes only 20.2% (17.3%) total power as compared to Si-CMOS. The EDP of the Ideal DG (SG) SB-GNRFET is only 1.0% (1.5%) of the MG Si-CMOS, while non-ideal DG (SG) SB-GNRFET has 133% (139%) EDP as compared to Si-CMOS. Non-ideal SB-GNRFET consumes less power than the ideal one because both its I_{on} and I_{off} are decreased by the presence of edge roughness. This decrease also reduces the I_{on}/I_{off} ratio, making the transistor less efficient, as can be seen in the degradation in delay and EDP.

C. Technology Scaling

We investigate the trend of delay and power when the transistor size scales down. The technology nodes available in the MG Si-CMOS PTM library are 16 nm, 14 nm, 10 nm, and 7 nm, with the supply voltage of 0.85 V, 0.8 V, 0.75 V, and 0.7 V, respectively. We scale the SB-GNRFET accordingly by putting in appropriate numbers of ribbons in one transistor. Fig. 9 shows the delay, dynamic power, leakage power, total power, and EDP of Si-CMOS, ideal and non-ideal DG and SG SB-GNRFET as the transistor size scales down. We show that delay, dynamic power, leakage power, total power, and EDP all scale down consistently across different technologies as the transistor size scales down, except for the leakage power of Si-CMOS, which increases with the downscaling. As a result, SB-GNRFETs show an advantage on the trend of power over Si-CMOS; Si-CMOS has almost constant total power with technology scaling, while the total power of SB-GNRFETs reduce as the transistor size scales down.

Moreover, ideal SB-GNRFETs give 1-2 orders of magnitude lower EDP than that of Si-CMOS. This indicates SB-GNRFET's potential in high-performance, low-power computing. Non-ideal SB-GNRFETs have reduced current and consume less power than the ideal ones, resulting in a very low total power. However, due to the significant increase in delay, the EDP of non-ideal SB-GNRFETs is much worse than that of the ideal ones. This brings up a pressing challenge for the fabrication technology to significantly improve the quality of GNR, especially investigating new techniques that can help produce GNRs with much smoother edges.

Nonetheless, an advantage of GNRFETs in terms of transistor size scaling is that they can scale based on the number of ribbons in one transistor. Therefore, only the driving strength of the transistor is scaled down with the transistor size, and the effect from scaling is linear. This results in stable and consistent circuit performance after scaling, which is also helpful in the circuit design process.

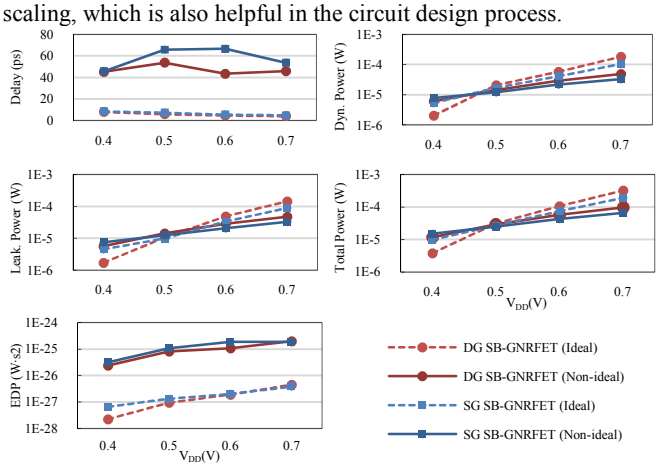

Fig. 7. Average delay, power, and EDP of benchmarks with V_{DD} scaling.

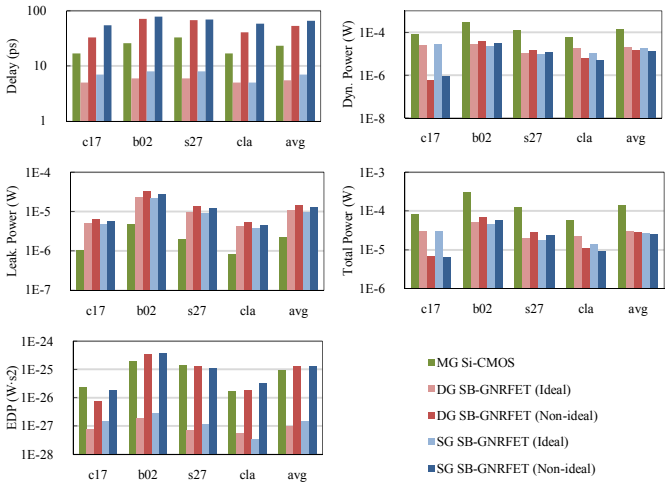

Fig. 8. Delay, power, and EDP of benchmark circuits.

VI. **Conclusion**

We present an analytical SPICE-compatible compact model of SB-GNRFETs that captures the effects of different parameters such as N (W_{CH}), L_{CH}, T_{ox}, T, and edge roughness. We analytically approximate both carrier charge density and carrier current in order to achieve closed-form expressions that make compact modeling of SB-GNRFETs possible. This model enables accurate and fast circuit simulation. Based on this model, we performed circuit-level performance evaluations on single-gate and double-gate SB-GNRFETs, with and without the impact of process variation. SB-GNRFET circuits are also compared with Si-CMOS-based circuits. We show that single-gate and double-gate SB-GNRFETs do not have a significant difference in delay, power, and EDP performance. SB-GNRFETs are also shown to perform better than Si-CMOS in terms of EDP. In the 16-nm node, the EDP is only ~1.3% of that of Si-CMOS for the ideal case, and ~136% for the non-ideal case. In the 7-nm node, the EDP is ~0.88% of that of Si-CMOS for the ideal case, and ~54% for the non-ideal case. These results indicate that the ideal SB-GNRFET has great performance and scalability, demonstrating its potential in becoming a next-generation device. However, advanced fabrication techniques are required to remove the non-idealities faced by GNR fabrication now before GNRFETs can become a competitive alternative solution beyond Si-CMOS.

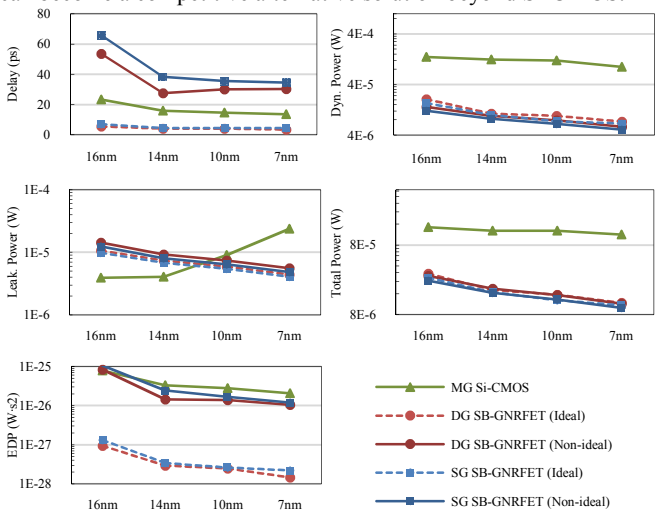

Fig. 9. Average delay, power, and EDP with technology scaling on the four benchmark circuits.

References

[1] C. Berger et al., "Electronic Confinement and Coherence in Patterned Epitaxial Graphene," *Science*, vol. 312, pp. 1191–1196, 2006.

[2] Z. Chen, Y.-M. Lin, M. J. Rooks, and P. Avouris, "Graphene Nano-ribbon Electronics," *Physica E: Low-dimensional Systems and Nanostructures*, pp. 228-232, 2007.

[3] A. K. Geim and K. S. Novoselov, 2007 *Nat. Mater*. 6 183.

[4] P. Wessely, F. Wessely, E. Birinci, U. Schwalke, and B. Riedinger, "Transfer-free Fabrication of Graphene Transistors," *J. Vac. Sci. Technol*. B 30, 2012.

[5] C. Sire, et al. "Flexible Gigahertz Transistors Derived from Solution-Based Single-Layer Graphene," *Nano Lett*., 2012.

[6] X. Wang, Y. Ouyang, X. Li, H. Wang, J. Guo, and H. Dai, "Room-Temperature All-Semiconducting Sub-10-nm Graphene Nanoribbon Field-Effect Transistors," *Phys. Rev. Lett*., 2008.

[7] L. Jiao, L. Zhang, L. Ding, and H. Dai, "Aligned Graphene Nanoribbons and Crossbars from Unzipped Carbon Nanotubes," *Nano Res*. 2010.

[8] S. Prezioso, F. Perrozzi, M. Donarelli, F. Bisti, S. Santucci, L. Palladino, M. Nardone, E. Treossi, V. Palermo, and L. Ottaviano, "Large Area Extreme-UV Lithography of Graphene Oxide via Spatially Resolved Photoreduction," *Langmuir*, 2012.

[9] G. Fiori and G. Iannaccone, "Simulation of Graphene Nanoribbon Field Effect Transistors," *Elec. Dev*. Lett., 2007.

[10] G. Fiori and G. Iannaccone, "Code for the 3D Simulation of Nanoscale Semiconductor Devices, Including Drift-diffusion and Ballistic Transport in 1D and 2D Subbands, and 3D Tunneling," *J. Comput. Electron*, vol. 4, no. 1/2, pp. 63–66, Apr. 2005.

[11] NanoTCAD ViDES. [Online]. http://vides.nanotcad.com/vides/

[12] P. Michetti, and G. Iannaccone, "Analytical Model of One-dimensional Carbon-based Schottky-barrier Transistors," *IEEE Trans. on Electron Devices*, vol. 57, no. 7, 2010.

[13] X. Yang, G. Fiori, G. Iannaccone, and K. Mohanram, "Semi-analytical Model for Schottky-barrier Carbon Nanotube and Graphene Nanoribbon Transistors," *GLSVLSI*, 2010.

[14] M. R. Choudhury, Y. Yoon, J. Guo, and K. Mohanram, "Graphene Nanoribbon FETs: Technology Exploration for Performance and Reliability," *Transactions on Nanotechnology*, 2011.

[15] Y.-Y. Chen, A. Sangai, M. Gholipour, and D. Chen, "Schottky-Barrier-Type Graphene Nano-Ribbon Field-Effect Transistors:A Study on Compact Modeling, Process Variation, and Circuit Performance," *NanoArch* 2013.

[16] S. Sinha, G. Yeric, V. Chandra, B. Cline, and Y. Cao, "Exploring Sub-20nm FinFET Design with Predictive Technology Models," *Design Automation Conference*, 2012.

[17] S. Vaziri, A.D. Smith, C. Henkel, M. Ostling, M.C. Lemme, "An Integration Approach for Graphene Double-Gate Transistors," *European Solid-State Device Research Conference*, 2012.

[18] S. Chilstedt, C. Dong, and D. Chen, "Carbon Nanomaterials Transistors and Circuits, Transistors: Types, Materials and Applications," *Nova Science Publishers*, 2010.

[19] M. Y. Han, B. Özyilmaz, Y. Zhang, and P. Kim, "Energy Band Gap Engineering of Graphene Nanoribbons," *Phys. Rev. Lett*. , vol. 98, no. 20, p. 206 805, May 2007.

[20] Y. W. Son, M. L. Cohen, and S. G. Louie, "Energy Gaps in Graphene Nanoribbons," *Phys. Rev. Lett*., vol. 97, no. 21, 216803, Nov. 2006.

[21] D. Jimenez, "A Current–voltage Model for Schottky-barrier Graphene-based Transistors," *Nanotechnology* 19, 2008.

[22] K. Nakada and M. Fujita, "Edge State in Graphene Ribbons: Nanometer Size Effect and Edge Shape Dependence," *Phys. Rev*. B, (Condensed Matter), 1996.

[23] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejon, "Tight-binding Description of Graphene," Phys. Rev. B, *Condens. Matter*, vol. 66, no. 3, p. 035 412, Jul. 2002.

[24] J. W. Mintmire and C. T. White, "Universal Density of States for Carbon Nanotubes," *Phys. Rev. Lett*., vol. 81, no. 12, pp. 2506–2509, Sep. 1998.

[25] N. H. Frank and L. A. Young, "Transmission of Electrons through Potential Barriers," *Phys. Rev*., 1931.

[26] FastCap 2. [Online]. http://www.fastfieldsolvers.com

[27] J. Deng and H.-S. Wong, "A Compact SPICE Model for Carbon-nanotube Field-effect Transistors Including Nonidealities and its Application Part I: Model of the Intrinsic Channel Region," *TED*, 2007.

[28] S. Datta, "Electronic Tansport in Mesoscopic Systems," *Cambridge University Press*, 1995.

[29] Z. Chen, J. Appenzeller, Y.-M. Lin, J. Sippel-Oakley, A. G. Rinzler, J. Tang, S. J. Wind, P. M. Solomon, and P. Avouris, "An Integrated Logic Circuit Assembled on a Single Carbon Nanotube," *Science*, 2006.

[30] [Online]. http://www.ptm.asu.edu