J Comput Electron (2009) 8: 441–450
DOI 10.1007/s10825-009-0282-2

An investigation of performance limits of conventional and tunneling graphene-based transistors

R. Grassi · A. Gnudi · E. Gnani · S. Reggiani · G. Baccarani

Published online: 26 August 2009
© Springer Science+Business Media LLC 2009

Abstract In this paper we perform a simulation study on the limits of graphene-nanoribbon field-effect transistors (GNR-FETs) for post-CMOS digital applications. Both conventional and tunneling FET architectures are considered. Simulations of conventional narrow GNR-FETs confirm the high potential of these devices, but highlight at the same time OFF-state leakage problems due to various tunneling mechanisms, which become more severe as the width is made larger and require a careful device optimization. Such OFF-state problems are partially solved by the tunneling FETs, which allow subthreshold slopes better than 60 mV/dec, at the price of a reduced ON-current. The importance of a very good control on edge roughness is highlighted by means of a direct simulation of devices with non-ideal edges.

Keywords Graphene nanoribbons · Carbon electronics · Nanoelectronic devices · Tunneling FET

1 Introduction

Graphene is a monolayer of carbon atoms arranged in a honeycomb lattice. Since its discovery in 2004 [1], it has emerged as one of the most promising alternatives to silicon for the fabrication of high-performance switching devices, due to its remarkable properties. Graphene samples have been experimentally produced through an exfoliation process of graphite [1] or by means of epitaxial growth [2]. The fundamental electronic properties of graphene are summarized hereafter:

- Graphene is a semimetal (or zero gap semiconductor): its energy dispersion relationship can be approximated by a conical surface defined in the 2D reciprocal lattice around the point of interception of the valence and conduction bands. Under charge neutrality conditions the Fermi level is at the interception energy between valence and conduction bands, but can be shifted with the application of a vertical electric field to create a majority of holes or electrons.
- Graphene exhibits very high carrier mobilities at room temperature due to a weak electron-phonon interaction. Mobilities as large as 10,000 cm^2/Vs have been measured with a very weak temperature dependence down to 4 K, indicating that impurity scattering was the limiting factor in those measurements [3]. Hence, electrons can travel ballistically in graphene over long distances (of the order of one micrometer) which by far exceed the length of advanced FETs.
- Graphene can sustain current densities exceeding those of copper at comparable dimensions.

Being graphene a 2D zero gap semiconductor, its use in nanoelectronic devices such as FETs requires a gap to be opened, which can be achieved by reducing the lateral size (nanoribbons), thus exploiting quantum confinement and the formation of subbands. Graphene nanoribbons (GNR) can be obtained either through lithographical etching (see, e.g., [4]), or through chemical processes [5], and the possibility of band gap engineering has been experimentally demonstrated [6]. Theoretical calculations predict that the energy dispersion relationships and intersubband gaps of ideal GNRs depend on the orientation and width. More

R. Grassi · A. Gnudi (✉) · E. Gnani · S. Reggiani · G. Baccarani
ARCES—Advanced Research Center for Electronics Systems, DEIS—Department of Electronics, University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy
e-mail: agnudi@arces.unibo.it

specifically, GNRs with armchair edges are always semiconducting with width-dependent gaps, while GNRs with zigzag edges are always nearly metallic, with sharply localized edge states near the Fermi level.

These important properties make GNRs potentially suitable for the fabrication of both n-type and p-type ultra-thin FETs, as well as of low-resistance interconnects on the same graphene sheet, thus retaining, and possibly improving, the circuit advantages of CMOS technology. In GNR FETs one can expect a very good control of short channel effects due to the monolayer thickness and an excellent switching speed due to the large carrier mobility and the long carrier mean-free path. GNR FETs with widths ranging from several tens of nanometers down to 2 nm have been fabricated and experimentally characterized [5]. The obtained results, however, are still far from satisfactory, due to a number of reasons mainly related with the difficulty of achieving a good edge control, leading to edge roughness and defects, and to the use of metallic source/drain regions, with the consequent formation of Schottky barriers, in order to avoid the still open problem of doping graphene layers.

In this paper we present a simulation study of both conventional and tunneling GNR-FET architectures. For the former architecture, after notifying the clear advantage over silicon of extremely narrow (1–2 nm) GNR-FETs, we address the problem of how and to what extent it is possible to relax the GNR width while maintaining a high performance level. More specifically, we investigate the design parameter space in order to cope with the small band gap of GNRs of controllable width, while maintaining a minimum ON/OFF current ratio. For the tunneling FET architecture, the investigation also provides some guidelines for the choice of the design parameters, showing the great potential for very good ON/OFF ratios at low supply voltages. The impact of edge roughness is finally assessed.

The paper is organized as follows. In the next section we briefly discuss the simulation models used in this work. Section 3 illustrates the simulation results for the ideal conventional GNR-FET, while Sect. 4 is devoted to the study of the tunneling GNR-FETs. Conclusions are drawn in Sect. 5.

2 Simulation models

Two different types of transport models are used in this paper: (i) an atomistic tight-binding (TB) model [7]; (ii) a non-parabolic effective-mass (NPEM) model [8]. In both cases, the 3D Poisson equation is solved self-consistently for the electrostatic potential. The TB model can be considered the state-of-the-art for quantum ballistic simulation of GNR-FETs, and is generally believed to accurately describe the transport physics in carbon-based devices [9], thus allowing for a deep physical insight. However, it is not well suited for repeated use in the design and optimization phases. For this purpose we have developed the NPEM model, a more agile approach based on the effective mass approximation. By introducing suitable non-parabolic corrections in the Hamiltonian via a position dependent effective mass, the method becomes accurate and compares well with TB even in the cases in which tunneling through the energy gap is important.

2.1 Tight-binding model

The adopted TB model is based on nearest-neighbor atom interaction, single p_z orbital and hopping energy correction to account for ribbon edge relaxation. Indicating with $|l,\alpha\rangle$ the orbital associated with the α-th atom within the l-th 1D primitive cell of the nanoribbon (slab), the matrix element of the Hamiltonian is written as

$$\langle l,\alpha|H|m,\beta\rangle = t_{l\alpha,m\beta} + \delta_{l\alpha,m\beta}U_{l\alpha}, \tag{1}$$

where $U_{l\alpha}$ is the potential energy at the (l,α) atom site, $\delta_{l\alpha,m\beta}$ is the Kronecker delta tensor, and $t_{l\alpha,m\beta}$ is equal to t if the atoms (l,α) and (m,β) are nearest neighbors and zero otherwise.

The hopping energy t is chosen by calibrating the TB energy dispersion relationship on those obtained from first principle calculations. In [10] it was found $t = 2.7$ eV for all atom pairs, except for those along the ribbon edges for which t should be incremented by the factor $(1+\delta)$ with $\delta = 0.12$. We checked the validity of the calibrated parameters against density functional theory (DFT) calculated subbands [11]. The first conduction and valence subband pairs for a $(N_a, 0)$ armchair GNR, where N_a is the number of atoms in the transverse zigzag cross-section of the GNR, are shown in Figs. 1 and 2 for $N_a = 6$ and $N_a = 13$, respectively. For reference, a (13,0) GNR slab is shown in Fig. 3. The DFT subbands have been calculated with hydrogen-saturated edge bonds with and without geometry relaxation, achieved by minimization of the total energy. It is seen that geometry relaxation is only responsible for minor differences and can be neglected. For both N_a values, the energy gaps and the shape of the first two subband pairs around the minima are well reproduced by the TB model. Some discrepancies are detected at higher energies, which are not so important for the device electrical characteristic.

The transport problem is formulated within the Non Equilibrium Green's Function (NEGF) formalism [12] based on the above Hamiltonian. The effect of the contacts is included via the self energies $\Sigma^r_{S/D}$, and the retarded Green's function is given by

$$G^r(E) = \left[EI - H - \Sigma^r_S - \Sigma^r_D\right]^{-1}. \tag{2}$$

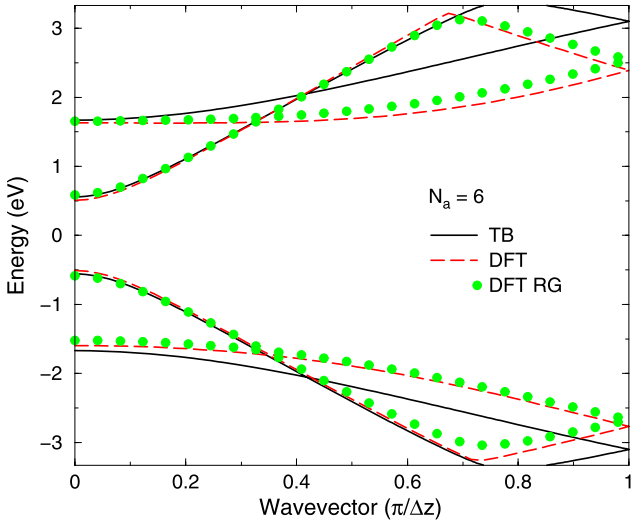

Fig. 1 Energy dispersion relationship vs. wavevector normalized to $\pi/\Delta z$, where $\Delta z = 3a_{cc}$ is the slab length with a_{cc} the interatomic distance, for an $(N_a, 0)$ GNR with $N_a = 6$. *DFT*: density functional theory calculations with the nominal geometry of graphene; *DFT RG*: density functional theory with relaxed geometry after using energy minimization; *TB*: tight binding model with parameters $t = 2.7$ eV and $\delta = 0.12$

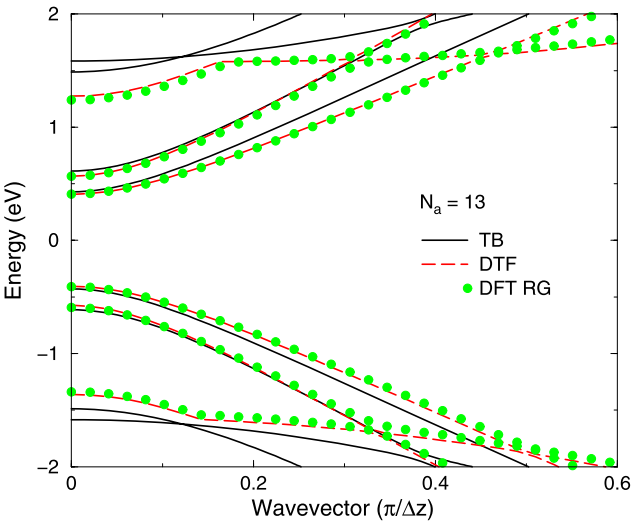

Fig. 2 Same as Fig. 1 for a $(13, 0)$ GNR

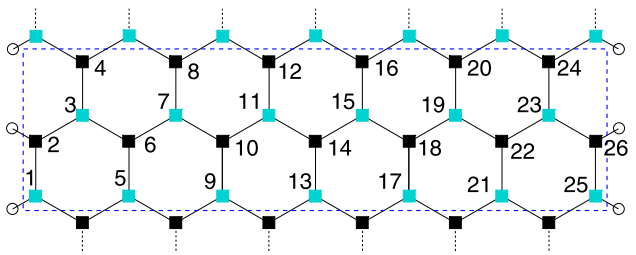

Fig. 3 Elementary slab of a $(13, 0)$ armchair GNR

The electron/hole correlation functions are given by $G^{</>} = G^r \Sigma^{</>} G^{r\dagger}$, from which the electron and hole numbers at the (l, α) atom site read

$$n_{l\alpha} = -2\mathrm{i} \int_{E_i(l,\alpha)}^{\infty} \frac{\mathrm{d}E}{2\pi} G^{<}(l, \alpha; l, \alpha; E),$$
$$p_{l\alpha} = 2\mathrm{i} \int_{-\infty}^{E_i(l,\alpha)} \frac{\mathrm{d}E}{2\pi} G^{>}(l, \alpha; l, \alpha; E), \tag{3}$$

where $E_i(l, \alpha)$ is the intrinsic Fermi level, assumed equal to the potential energy $U_{l\alpha}$.

Indicating with $G^{</>/r}(l, m; E)$ the submatrix relative to the (l, m) pair of slabs for energy E, the current is calculated as

$$I = \frac{q^2}{h} \int_{-\infty}^{\infty} \mathrm{d}E \times \frac{4}{q} \Re \left\{ \mathrm{Tr} \left[H(l, l+1) G^{<}(l+1, l; E) \right] \right\}, \tag{4}$$

where the symbols $\Re$ and Tr indicate the real part and the trace on the orbital index, respectively.

The electrostatic potential is calculated by self consistently solving the 3D Poisson equation. The box integration method is used on a discretization grid of prismatic elements with a triangular base, matching the hexagonal graphene lattice. The electron and hole charge given by (3) is directly assigned to the box surrounding the (l, α) atom.

2.2 Effective mass with non-parabolic corrections

In this case the full 2D quantum problem is separated into the transverse (lateral confinement) and longitudinal (transport) directions. In principle, the energy dispersion relationships (subbands) $\varepsilon(k)$ should be calculated for every slab l of the ribbon starting from the TB Hamiltonian with the self-consistent potential energy U repeated periodically throughout the device. As an example, Figs. 4 and 5 report the resulting $\varepsilon(k)$, assuming a potential energy identically equal to zero, for the first two subbands (green symbols) of the $(6, 0)$ and $(13, 0)$ GNRs previously considered. The portions of the dispersion relationships with energies in the gaps and purely-imaginary k, corresponding to vanishing states [13], are included in the figures. Also shown are the dispersion relationships (solid lines) obtained with the non-parabolic expression

$$\left(\varepsilon_b - \frac{E_g^b}{2} \right) \left(\frac{1}{2} + \frac{\varepsilon_b}{E_g^b} \right) = \frac{\hbar^2 k^2}{2 m_b^\star}, \tag{5}$$

where b is the subband index, $E^b{}_g$ is the energy gap and $m_b^\star$ is the electron effective mass in the same subband, which is treated as a fitting parameter. The two sets of curves are

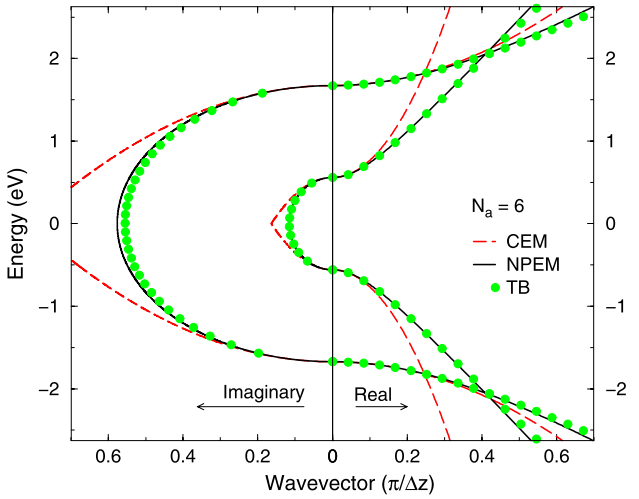

Fig. 4 Energy dispersion relationships for the two lowest conduction/valence subband pairs of a (6, 0) GNR calculated with the TB model, the effective mass model with non-parabolic corrections (NPEM) and the constant effective mass (CEM) model, as a function of the normalized wavevector. Both real and imaginary wavevectors are considered

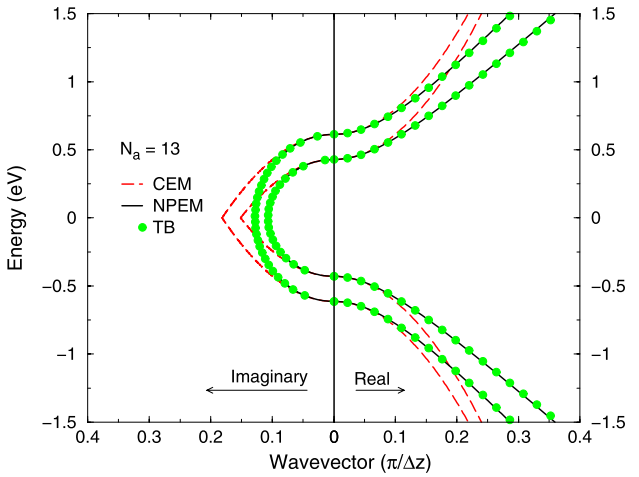

Fig. 5 Same as in Fig. 4 but for a (13, 0) GNR

in excellent agreement over an extended range of energies, including those in the gap. To show the importance of the non-parabolic corrections, Figs. 4 and 5 report the parabolic dispersion relationships obtained by setting the second factor in (5) equal to one. It should be noticed that the main problem with this constant effective-mass (CEM) model occurs for energies in the gap, leading to an inaccurate estimation of the tunneling currents.

Very similar subband shapes are obtained with non-zero potential energies, suggesting the possibility of neglecting the differences of the energy dispersion relationships relative to the minima originating from the different device slabs. A further simplification arises from neglecting the k-dependence of the transverse wavefunctions. The above

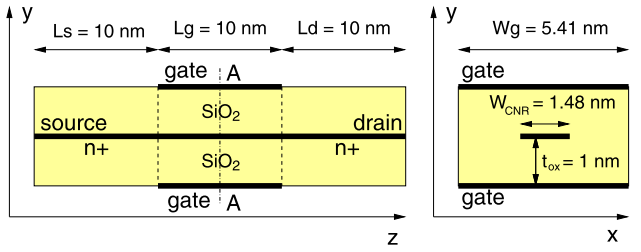

Fig. 6 Longitudinal (*left*) and transverse (*right*, section AA) cross-section of the simulated double-gate (13, 0) GNR-FET

considerations lead to a simplified solution of the transverse problem: first, for a given N_a the full subband structure is computed only once for a slab at zero potential in order to extract the effective masses of the lowest subbands to be used in the successive transport calculations; next, the eigenvalues are computed with the TB model for every slab only at $k = 0$, so as to obtain the bottom (top) of the conduction (valence) bands, as well as the corresponding eigenfunctions, which turn out to be parametrically dependent on the longitudinal coordinate.

The transport problem is treated within the NEGF formalism, by solving a 1D transport equation in the longitudinal direction for every pair of conduction/valence subbands. The 1D transport Hamiltonian is written within the EM approximation, using the eigenvalues computed slab by slab as the electron/hole potential-energy profiles and the position/energy dependent effective mass derived from (5) to account for the non-parabolic behavior [8]. The computational advantage with respect to the TB approach is quite large.

To demonstrate the effectiveness of the NPEM model, we have simulated a (13, 0) and a (6, 0) double-gate GNR-FET. The geometry of the (13, 0) device is depicted in Fig. 6. The (6, 0) GNR-FET is similar, but with a width of 0.7 nm. The other parameters are reported in the figure. The gate-aligned source and drain regions are doped with an uniform effective molar fraction equal to 10^{-2}, while the channel is intrinsic. A remark is necessary at this point on the validity of the assumption of uniformly distributed doping concentration. In experimental carbon nanotube devices, heavy doping concentrations have been generated either chemically or electrostatically [14], and the same can be assumed for GNRs. In case of chemical doping, due to the very small number of carbon atoms in the source and drain regions, the assumption of average doping concentration is clearly an idealization, whose consequences should be carefully analyzed depending also on the device architecture. The main reason for its widespread use is to avoid the computational complexity of a statistical analysis carried out with respect to the position of the doping atoms. In case of electrostatic doping, the doping level must be interpreted as an effective value which contributes to fixing the electrostatic potential within source

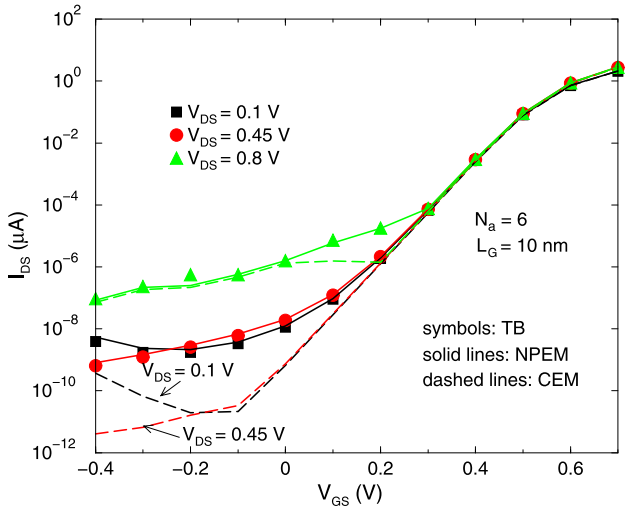

Fig. 7 Turn-on characteristics of the $(6, 0)$ GNR-FET computed at different V_{DS} with the tight-binding (TB), non-parabolic effective mass (NPEM) and constant effective mass (CEM) models

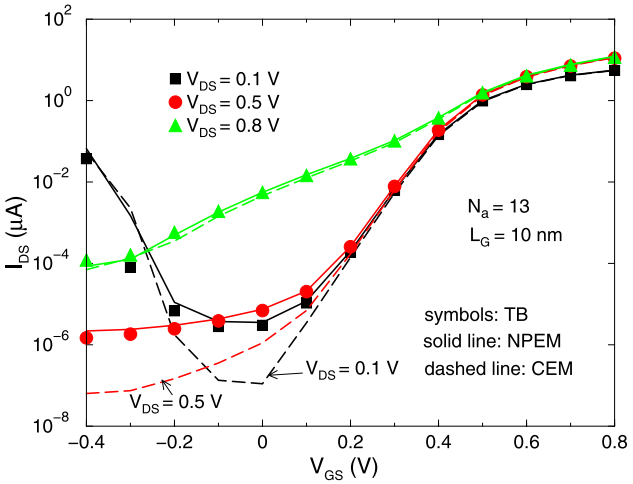

Fig. 8 Same as in Fig. 7 but for the $(13, 0)$ GNR-FET

and drain regions, therefore the assumption of uniformity is justified. The turn-on characteristics of the two GNR-FETs, computed with the TB, the NPEM and the CEM models for different V_{DS} values, are reported in Figs. 7 and 8. It can be seen that the NPEM model agrees remarkably well with the TB model, while the CEM model suffers from apparent limitations in certain bias regions, especially at low V_{GS}. Indeed, when the device operates under deep subthreshold conditions at the given gate length, transport is dominated either by direct tunneling (DT) from source to drain, or by band-to-band-tunneling (BTBT), which is responsible for the current rise at negative V_{GS} clearly visible at $V_{DS} = 0.1$ V in the $(13, 0)$ FET. One can therefore draw the conclusion that the simple CEM model lacks accuracy, especially in the bias regions where the DT current dominates. This can be ascribed to the poor description of the energy dispersion relationship in the gap evidenced in Figs. 4 and 5. The introduction of non-parabolic corrections with the NPEM model greatly improves the picture.

3 Simulation study of conventional GNR-FETs

From the turn-on characteristics shown above, the great potential of ideal and extremely-narrow graphene-based FETs can be fully appreciated. For example, from Fig. 8 a current drive capability normalized to the GNR width ($W = 1.5$ nm) as large as 8 mA/μm (at $V_{DS} = V_{GS} = 0.8$ V) is predicted, which by far exceeds that of present silicon devices at comparable supply voltages. On the other hand, the device suffers from limitations in turning the current off at the largest drain bias (Fig. 8, $V_{DS} = 0.8$ V), due to the onset of BTBT effects at the drain end of the channel, that reduce the gate control over the channel. Due to the formidable technological challenge in building GNRs of nanometer widths with well controlled edges, it is interesting to investigate how much the GNR width can be relaxed while maintaining an overall competitive device performance. Relaxing the GNR width means reducing the band gap; hence, the OFF-state current limitations are expected to become more severe, thus requiring a careful choice of the design parameters. In this section we present a performance investigation of relatively wide (a few nanometers) GNR-FETs with small band gap using the NPEM model. Being the I_{ON}/I_{OFF} ratio the main limitation of small band gap devices, we discuss how the design parameters ought to be chosen in order to mitigate the problem. The ratio $I_{ON}/I_{OFF} > 10^4$ is chosen as the acceptance criterion. The topology is the same as in Fig. 6 with doped source/drain regions; however two intrinsic source/drain gate underlap regions of extension L_U have also been considered, where L_U is an optimization parameter. The gate length is fixed at $L_G = 20$ nm, so as to prevent direct source-to-drain tunneling in the OFF state.

The turn-on characteristics of a $(40, 0)$ 4.8 nm-wide GNR-FET are shown in Fig. 9 for $V_{DS} = 0.1$, 0.3 and 0.4 V with $L_U = 0$ and $L_U = 20$ nm. Here $\varepsilon_{ox} = 3.9$, $t_{ox} = 1$ nm, and the dopant molar fraction in the source and drain regions is equal to 10^{-3}. Clearly, for the self aligned device ($L_U = 0$) the I_{ON}/I_{OFF} ratio rapidly deteriorates for $V_{DS} > 0.3$ V. This can be understood by looking at the conduction and valence band profiles of Fig. 10 for $V_{DS} = 0.4$ V and $V_{GS} = 0.1$ V. As the band gap is only 0.29 eV, channel-to-drain BTBT occurs in the OFF state. The underlap regions make the potential profile at the drain junction smoother and effectively reduce I_{OFF}, but degrade at the same time I_{ON}. Figure 11 illustrates how the intrinsic underlap region at the source side creates a potential barrier

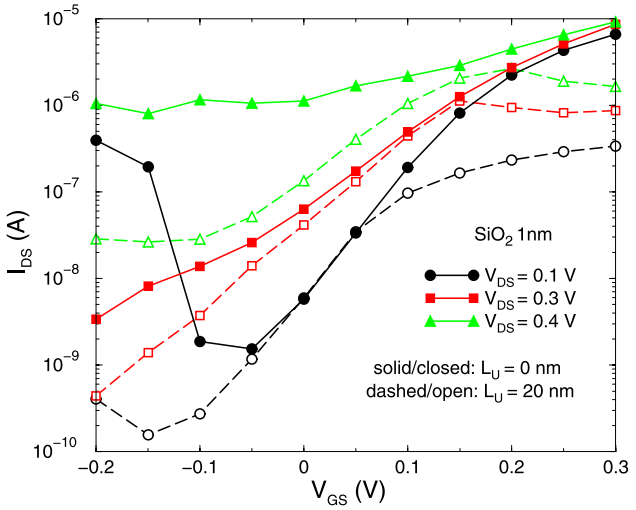

Fig. 9 Turn-on characteristics of the (40, 0) GNR-FET with $\varepsilon_{ox} = 3.9$ and $t_{ox} = 1$ nm at various V_{DS}, with ($L_U = 20$ nm) and without ($L_U = 0$) underlap

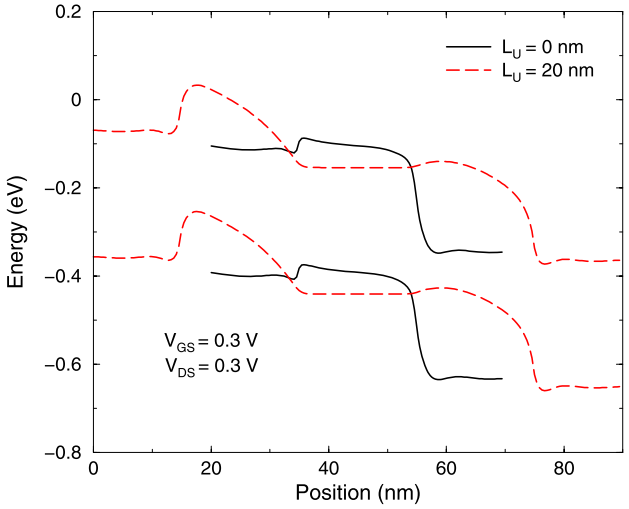

Fig. 11 Conduction and valence band profiles for the GNR-FET of Fig. 9 at $V_{GS} = V_{DS} = 0.3$ V with ($L_U = 20$ nm) and without ($L_U = 0$) underlap

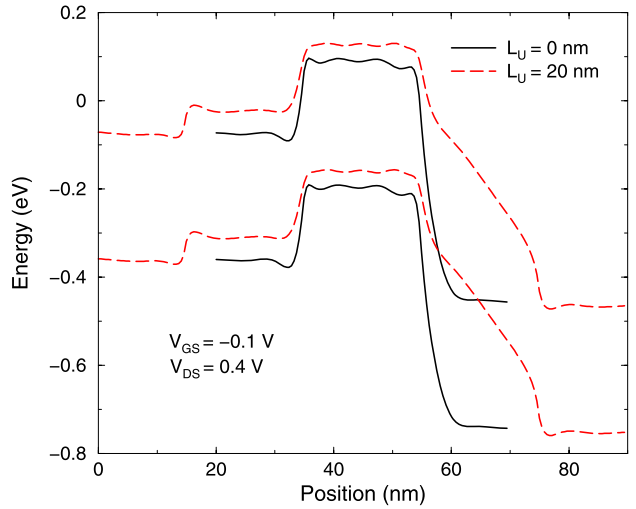

Fig. 10 Conduction and valence band profiles for the GNR-FET of Fig. 9 at $V_{GS} = 0.1$ V and $V_{DS} = 0.4$ V with ($L_U = 20$ nm) and without ($L_U = 0$) underlap

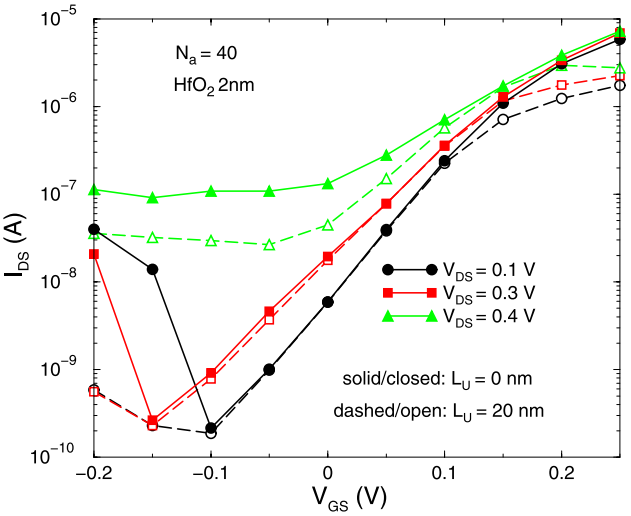

Fig. 12 Turn-on characteristics of the (40, 0) GNR-FET with $\varepsilon_{ox} = 16$ and $t_{ox} = 2$ nm at various V_{DS}, with ($L_U = 20$ nm) and without ($L_U = 0$) underlap

in the ON state that limits the peak current and is not controlled by the gate.

The effect of a high-κ dielectric is investigated next. The turn-on characteristics with $\varepsilon_{ox} = 16$ (HfO_2) and $t_{ox} = 2$ nm are reported in Fig. 12 for the same values of V_{DS}, both with and without underlap. From the comparison with Fig. 9, it appears that the increase of ε_{ox} has a beneficial effect mainly at low V_{GS}, considerably lowering the OFF current. At $V_{DS} = 0.4$ V the use of the underlap regions further reduces I_{OFF} which, however, remains unacceptably high. At $V_{DS} = 0.1$ and 0.3 V, the underlap regions simply deteriorate I_{ON} without any appreciable benefit on I_{OFF}. The use of underlap regions can therefore be ruled out.

From the previous analysis, we can argue that a ratio $I_{ON}/I_{OFF} = 10^4$ cannot be achieved by the devices considered so far. In order to reach the target ratio, we need a device able to bear $V_{DS} \simeq 0.4$ V without any significant degradation of the OFF current. Extrapolating from the previous considerations, we turn to a (28, 0) 3.3 nm-wide GNR-FET, having $E_G = 0.41$ eV. The turn-on characteristics, simulated with $\varepsilon_{ox} = 16$, $t_{ox} = 2$ nm, a source/drain dopant molar fraction of $1.5 \cdot 10^{-3}$ and $L_U = 0$ are shown in Fig. 13, while Fig. 14 illustrates the I_{ON}/I_{OFF} ratio at 0.3 and 0.4 V supply voltages for different I_{OFF} values, i.e. different gate work functions. For this device, the maximum achievable current ratio is nearly 10^5 for a supply voltage of

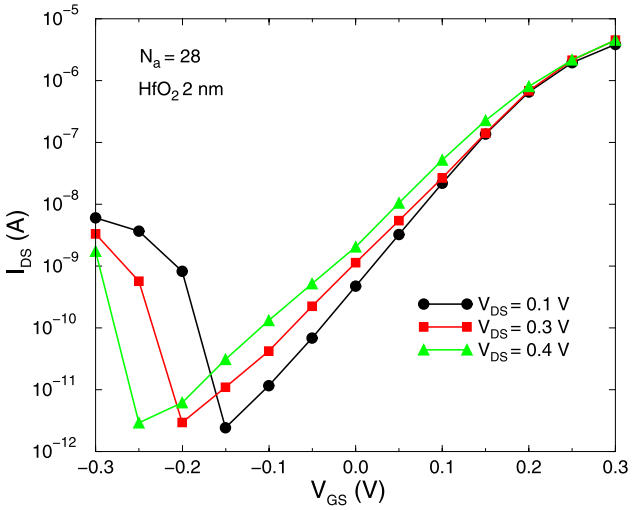

Fig. 13 Turn-on characteristics of the $(28, 0)$ GNR-FET with $\varepsilon_{ox} = 16$ and $t_{ox} = 2$ nm at various V_{DS} without underlap

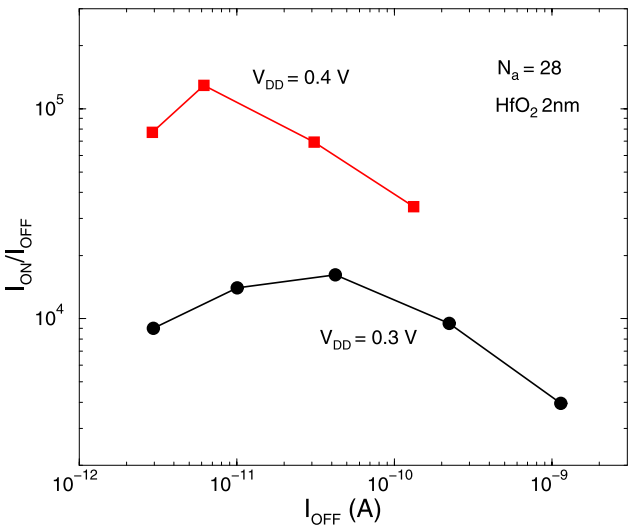

Fig. 14 I_{ON}/I_{OFF} ratio as a function of I_{OFF} (different gate work functions) for the GNR-FET of Fig. 13. Two supply voltages of 0.3 V and 0.4 V are considered

0.4 V. The current drive capability is 1.3 mA/μm (Fig. 13 at $V_{GS} = 0.3$ V and $V_{DS} = 0.4$ V) which is comparable with what obtained from silicon devices at 1 V supply. As the dynamic power scales with the square of the supply voltage, this lowering represents a nearly 6× advantage with respect to silicon.

In conclusion of this section we can state that, even if the best performance is obtained from nanometer-wide GNR-FETs, the width can be somewhat relaxed up to $\simeq 3.5$ nm while maintaining an acceptable ON/OFF current ratio in excess of 10^4. The resulting current drive capability for an ideal GNR-FET is comparable with that of silicon devices, with a definite advantage in terms of power dissipation.

4 Simulation study of tunneling GNR-FETs

The tunneling transistor (TFET) has been proposed as the ideal architecture for carbon nanotube (CNT) FETs [15], capable of overcoming some limitations of the conventional FET topology. The CNT-TFET is based on an intrinsic gated channel and source/drain regions with opposite types of doping, i.e., *n-i-p* or *p-i-n*. The band-bending in the source-channel junction is responsible for gate-controlled BTBT current which is the main conduction mechanism, as opposed to thermal emission over the barrier for conventional FETs. Thus, a subthreshold slope (SS) better than 60 mV/dec can be achieved. A similar behavior is expected from GNR-TFETs. In this section a number of GNR-TFETs are simulated and guidelines for the optimal choice of the design parameters (gate topology, type and size of dielectric, source/drain doping concentrations) are provided, reaching similar conclusions as in [16]. Besides, the effect of edge roughness is taken into account via the direct simulation of rough GNR channels with randomly generated defects at the edges. Indeed, edge roughness is known to seriously limit the performance of conventional GNR-FETs [5], reducing the ON current and increasing the OFF leakage, due to the formation of localized states in the gap [9].

We start with investigating the impact of some design parameters on the performance of ideal GNR-TFETs, with the purpose of understanding the key optimization issues. Two types of topologies are considered: the double gate (DG) topology, similar to Fig. 6 but with p^+ source, and the single gate (SG) one, similar to the former but with one top gate and a 10 nm thick bottom dielectric. In all simulated devices $W_G - W_{GNR} = 4$ nm.

The following general guidelines for the design of TFETs are known from the literature. In order to increase the ON current, the BTBT at the source junction must be favored, which suggests the use of GNRs with small band gap and the opportunity of having a high longitudinal electric field, i.e. band bending, at the same junction. In an ideal armchair GNR the band gap depends essentially on the ribbon width, while the shape of the electric field involves a number of parameters, such as the dielectric constant of the insulator, its thickness and the source doping concentration. On the other hand, the minimum leakage current for a given V_{DS} can be traced back to the three following mechanisms, which have different importance depending on the device parameters: thermal current (injection of conduction band electrons from the source and valence band holes from the drain), source-to-drain BTBT throughout the entire channel, and BTBT at the drain junction (responsible for current rise at low V_{GS}, similar to conventional FETs). All three conduction mechanisms can be reduced by making the band gap larger, in contrast with the large ON current requirement. Moreover, the condition for which the leakage current is given by BTBT at

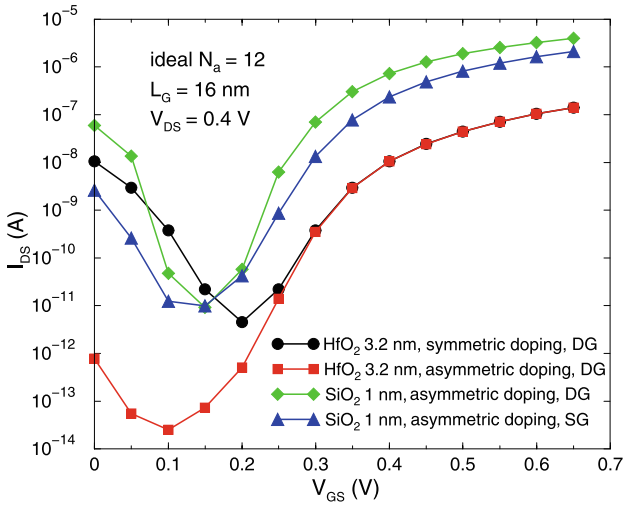

Fig. 15 Turn-on characteristics of the ideal $(12, 0)$ GNR-TFETs with $L_G = 16$ nm, source doping molar fraction $N_S = 5 \cdot 10^{-3}$ and $V_{DS} = 0.4$ V. Legends: $\varepsilon_{ox} = 16$ for HfO_2, $\varepsilon_{ox} = 3.9$ for SiO_2, $N_D = N_S$ for sym. dop., $N_D = 10^{-3}$ for asym. dop., double (DG) or single (SG) gate topology

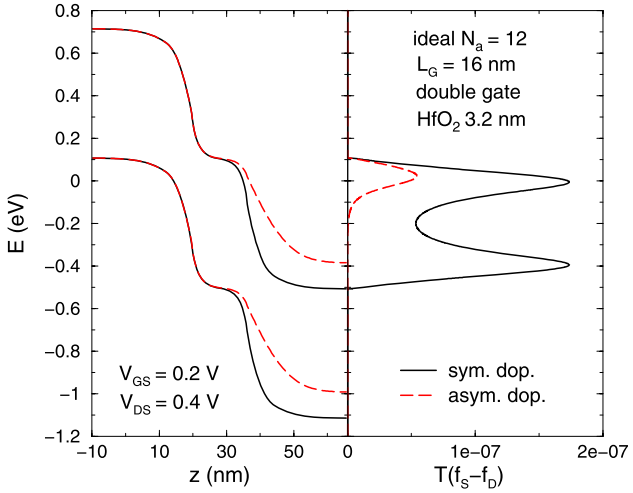

Fig. 16 Band diagram (*left*) and transmission coefficient multiplied by the Fermi function difference between source and drain (normalized current density) vs. energy (*right*) for the $(12, 0)$ ideal GNR-TFET with $V_{GS} = 0.2$ V and $V_{DS} = 0.4$ V for different doping levels. Legends are as in Fig. 15

the drain junction can be avoided by making the maximum V_{DS} sufficiently smaller than the gap.

The turn-on characteristics of different devices have been simulated. The results for an ideal $(12, 0)$ GNR ($W_{GNR} = 1.35$ nm, $E_G = 0.61$ eV) are reported in Fig. 15. The applied drain voltage $V_{DS} = 0.4$ V is sufficiently lower than E_G to ensure that the leakage current is not due to BTBT at the drain junction. First of all, we investigate the choice of the doping levels in the source and drain regions. The curves marked with black circles and red squares in Fig. 15 have been obtained with the same parameter set (please refer to the figure caption for details) except for the doping molar fraction in the drain, which is $5 \cdot 10^{-3}$ and 10^{-3}, respectively. The leakage current is clearly reduced for the lower drain doping level. The reason is better understood by observing the band diagram and current density plots of Fig. 16 relative to $V_{GS} = 0.2$ V. The current is mainly due to BTBT into the channel, and the effect of the lower doping concentration is twofold: (i) reducing the drain degeneracy by shifting up the conduction band edge and, (ii) making the potential transition from channel to drain less abrupt, thus elongating the tunneling path. Both effects reduce the BTBT current.

Next, we examine the effect of dielectric type and thickness. To this purpose, the red squares and green diamonds curves of Fig. 15 must be compared. The former is obtained with 3.2 nm HfO_2 ($\varepsilon_{ox} = 16$); the latter with 1 nm SiO_2 ($\varepsilon_{ox} = 3.9$), all other parameters being the same. Despite the larger EOT and the lower C_{OX}, the SiO_2 TFET exhibits strikingly larger currents. The band diagram and transmission coefficient plots of Fig. 17 for $V_{GS} = 0.55$ V reveal the importance of gate-source fringing. A thicker high-κ oxide

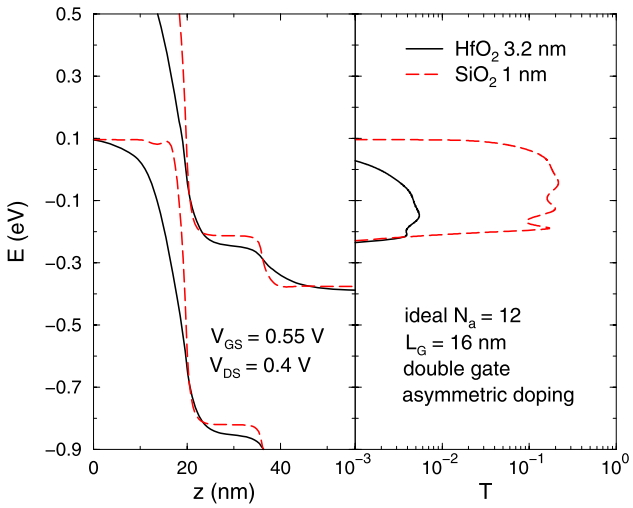

Fig. 17 Band diagram (*left*) and transmission coefficient vs. energy (*right*) for the $(12, 0)$ ideal GNR-TFET with $V_{GS} = 0.55$ V and $V_{DS} = 0.4$ V for various types of dielectrics. Legends as in Fig. 15

increases the fringing and reduces the BTBT at the source. It is also seen that the transmission coefficient is quite large ($T_{max} \simeq 0.3$), and that the BTBT does not represent a serious bottleneck for achieving high ON currents.

We also check the effect of removing the bottom gate by comparing SG (Fig. 15, blue triangles) and DG (green diamonds) topologies with the same parameters. It turns out that the DG TFET is preferable, since it leads to higher ON-currents and better SS. If we look at the performance of such devices, we obtain 1.89 mA/μm ON-current (Fig. 15, green diamonds, $V_{GS} = 0.55$ V) with an ON/OFF current ratio larger than 10^5. Hence, as compared to conventional FETs

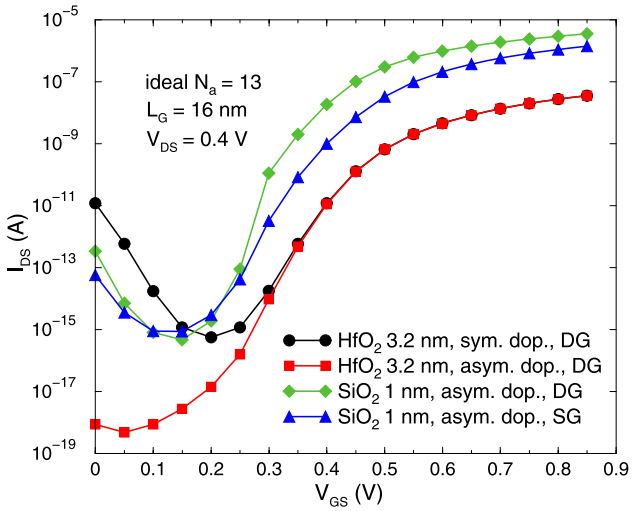

Fig. 18 Same as in Fig. 15 but for a (13, 0) GNR-TFET

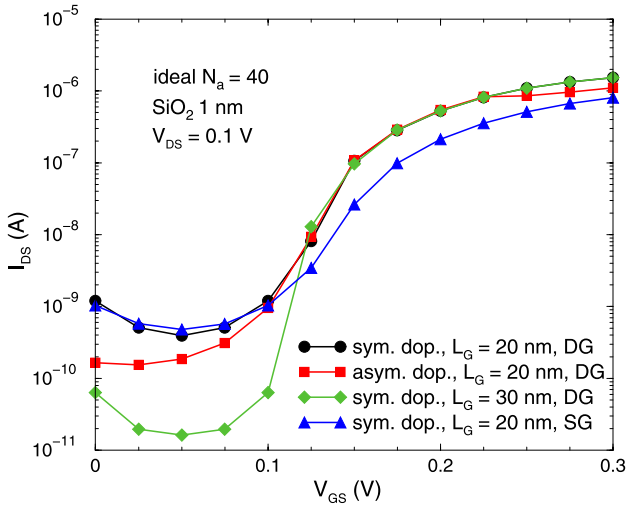

Fig. 19 Turn-on characteristics of the ideal (40, 0) GNR-TFETs with source doping molar fraction $N_S = 7 \cdot 10^{-4}$, 1 nm SiO_2 and $V_{DS} = 0.1$ V. Legends: as in Fig. 15, except $N_D = 2 \cdot 10^{-4}$ for asym. dop. and the use of different gate lengths

of similar width, TFETs easily provide large ON/OFF current ratios at the price of lower ON currents. Similar remarks apply to TFETs of different widths, as confirmed by the turn-on characteristics of (13, 0) and (40, 0) GNR-TFETs shown in Figs. 18 and 19, respectively. It should be noticed that the (13, 0) TFET exhibits an ON/OFF current ratio of 10^9 at $V_{DS} = 0.4$ V which exceeds that of its (12, 0) counterpart due to the slightly-larger band gap (Fig. 18, green diamonds). As far as the (40, 0) GNR-TFET is concerned ($W_{GNR} = 4.8$ nm, $E_G = 0.29$ eV), the performance is in general rather poor due to the quite small band gap which forces the maximum $V_{DS} = 0.1$ V in order to avoid BTBT at the drain. Only the best configuration with 1 nm SiO_2 is considered for this device. Current leakage is dominated

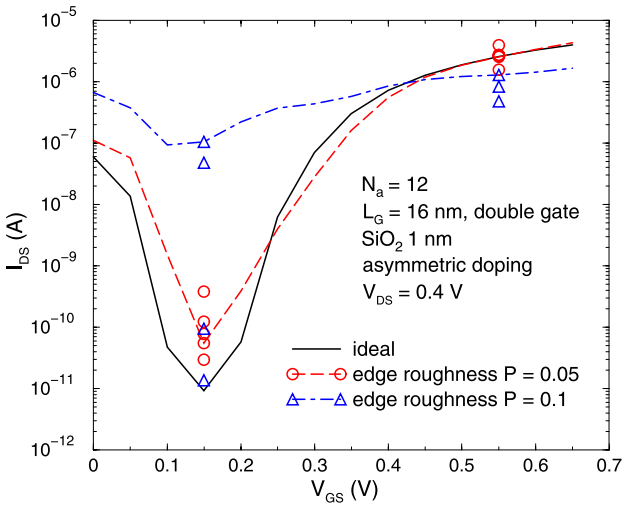

Fig. 20 Turn-on characteristics of nominal (12, 0) GNR-TFETs with edge roughness with different probabilities P (the ideal GNR corresponds to $P = 0$)

by BTBT through the channel. Increasing the gate length L_G helps suppress the leakage, as confirmed by the curve at $L_G = 30$ nm in Fig. 19. It is interesting to notice that, despite the very low V_{DS}, a remarkable ON/OFF ratio larger than 10^4 can be achieved, indicating the great potential of graphene for low dynamic power applications.

Finally, we examine the effect of edge roughness for the case of the best performing nominal (12, 0) GNR-TFET previously considered (Fig. 15, green diamonds). The edge defects are simulated by randomly adding or removing atom pairs at the two edges independently according to a predefined probability P, according to the approach proposed in [17]. The turn-on characteristics for $P = 0$ (ideal case), $P = 0.05$ and $P = 0.1$ are plotted in Fig. 20. It is seen that a moderate amount of defects can be tolerated, even if both the ON and OFF currents are deteriorated. However, a further increase of roughness can lead to the impossibility of turning the device off. The reason for this can be traced back to the onset of states in the gap, which increase the BTBT through the channel, as shown by the local density of states in Fig. 21. Moreover, different implementations of the edge defects with the same probability P may lead to different results for the OFF and ON currents, as reported in Fig. 20. This effect proves a critical variability problem.

5 Conclusions

A simulation study of both conventional and tunneling GNR-FETs has been presented, based on a quantum ballistic-modeling approach with either an atomistic TB or an effective mass approximation with non-parabolic corrections. The importance of such corrections is shown to be critical for the achievement of accurate tunneling currents.

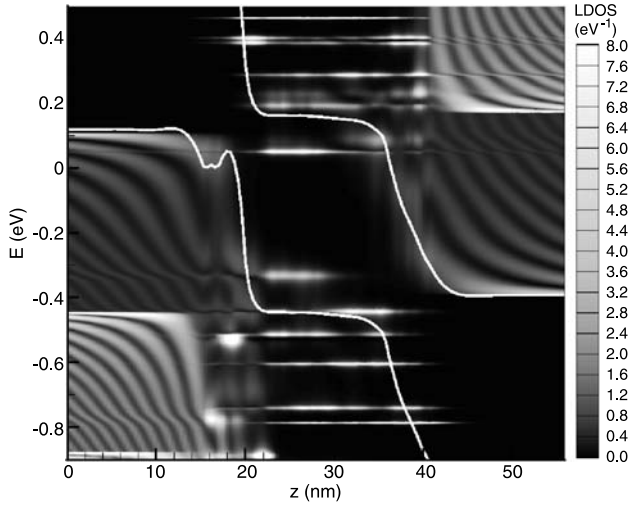

Fig. 21 Local density of states (LDOS) integrated on each slab of the (12, 0) rough GNR-TFET with $P = 0.1$, $V_{DS} = 0.4$ V, $V_{GS} = 0.15$ V

Simulation results indicate that extremely-narrow GNR-FETs with conventional architecture ($W \simeq 1.5$ nm) have a great potential for high-performance nanoelectronic devices, provided they are defect free and not affected by edge states. In this highly-idealized case, the ON current density is as large as 8 mA/μm, which by far exceeds the best performance of Si-based transistors with a comparable supply voltage. On the other hand, GNR devices suffer a limitation in the maximum allowable supply voltage, due to the BTBT which occurs at the drain end of the channel and severely degrades the OFF leakage current. Thus, GNR-FETs are especially suitable for low-voltage operation, at the expense of the ON/OFF current ratio, in line with requirements for high performance logic. If we relax the strict specification on the width, the smaller energy gap limits the maximum allowable supply voltage even further. In order to achieve a minimum ON/OFF current ratio equal to 10^4, the maximum GNR width ought to be around 3.3 nm. In this case, with 0.4 V supply voltage, ON currents comparable to silicon devices with 1 V supply are obtained, which still makes GNR-FETs very interesting for low dynamic power applications.

Some of the limitations of the conventional GNR-FETs can be removed with the TFET configuration, given the possibility of achieving an SS much better than 60 mV/dec. Some guidelines for the optimization of ideal GNR-TFETs have been provided: with the proper choice of design parameters an ON/OFF ratio larger than 10^9 is predicted for a (13, 0) TFET, at the expense of a somewhat reduced ON current capability. Thus, GNR-TFETs are possible candidates for trading off high performance and low power operation. On the other hand edge roughness, unless well controlled, spoils the device performance, in particular in the OFF state, and is also responsible for a large variability problem.

Acknowledgements This work has been supported by the EU FP7 IST Project GRAND (Contract n. 215752) via the IU.NET Consortium.

References

1. Novoselov, K.S., Geim, A.K., Morozov, S.V., Jiang, D., Zhang, Y., Dubonos, S.V., Grigorieva, I.V., Firsov, A.A.: Electric field effect in atomically thin carbon films. Science **306**, 666 (2004)
2. Berger, C., Song, Z., Li, X., Wu, X., Brown, N., Naud, C., Mayou, D., Li, T., Hass, J., Marchenkov, A.N., Conrad, E.H., First, P.N., de Heer, W.A.: Electronic confinement and coherence in patterned epitaxial graphene. Science **312**, 1191 (2006)
3. Geim, A.K., Novoselov, K.S.: The rise of graphene. Nat. Mater. **6**, 183 (2007)
4. Lemme, M.C., Echtermeyer, T.J., Baus, M., Kurz, H.: A graphene field-effect device. IEEE Electron Device Lett. **28**, 282 (2007)
5. Wang, X., Ouyang, Y., Li, X., Wang, H., Guo, J., Dai, H.: Room-temperature all-semiconducting sub-10-nm graphene nanoribbon field-effect transistors. Phys. Rev. Lett. **100**, 206803 (2008)
6. Han, M.Y., Özyilmaz, B., Zhang, Y., Kim, P.: Energy band-gap engineering of graphene nanoribbons. Phys. Rev. Lett. **98**, 206805 (2007)
7. Liang, G., Neophytou, N., Nikonov, D.E., Lundstrom, M.S.: Performance projections for ballistic graphene nanoribbon field-effect transistors. IEEE Trans. Electron Devices **54**, 677 (2007)
8. Grassi, R., Poli, S., Gnani, E., Gnudi, A., Reggiani, S., Baccarani, G.: Tight-binding and effective mass modeling of armchair graphene nanoribbon FETs. Solid-State Electron. **53**, 462 (2009)
9. Yoon, Y., Fiori, G., Hong, S., Iannaccone, G., Guo, J.: Performance comparison of graphene nanoribbon FETs with Schottky contacts and doped reservoirs. IEEE Trans. Electron Devices **55**, 2314 (2008)
10. Son, Y., Cohen, M.L., Louie, S.G.: Energy gaps in graphene nanoribbons. Phys. Rev. Lett. **97**, 216803 (2006)
11. Kohn, W., Sham, L.J.: Self-consistent equations including exchange and correlation effects. Phys. Rev. **140**, A1133 (1965)
12. Lake, R., Klimeck, G., Bowen, R.C., Jovanovic, D.: Single and multiband modeling of quantum electron transport through layered semiconductor devices. J. Appl. Phys. **81**, 7845 (1997)
13. Boykin, T.B.: Generalized eigenproblem method for surface and interface states: The complex bands of GaAs and AlAs. Phys. Rev. B **54**, 8107 (1996)
14. Lin, Y.-M., Appenzeller, J., Knoch, J., Avouris, Ph.: High-performance carbon nanotube field-effect transistor with tunable polarities. IEEE Trans. Nanotechnol. **40**, 481 (2005)
15. Appenzeller, J., Lin, Y.-M., Knoch, J., Chen, Z., Avouris, P.: Comparing carbon nanotube transistors—The ideal choice: A novel tunneling device design. IEEE Trans. Electron Devices **52**, 2568 (2005)
16. Zaho, P., Chauhan, J., Guo, J.: Computational study of tunneling transistor based on graphene nanoribbon. Nano Lett. **9**, 684 (2009)
17. Yoon, Y., Guo, J.: Effect of edge roughness in graphene nanoribbon transistors. Appl. Phys. Lett. **91**, 073103 (2007)