A Three-Dimensional Simulation Study of the Performance of Carbon Nanotube Field-Effect Transistors with Doped Reservoirs and Realistic Geometry

Gianluca Fiori
Giuseppe Iannaccone
Gerhard Klimeck

Purdue University - Main Campus, gekco@purdue.edu

Follow this and additional works at: http://docs.lib.purdue.edu/nanodocs

http://docs.lib.purdue.edu/nanodocs/111

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
A Three-Dimensional Simulation Study of the Performance of Carbon Nanotube Field-Effect Transistors With Doped Reservoirs and Realistic Geometry

Gianluca Fiori, Giuseppe Iannaccone, Member, IEEE, and Gerhard Klimeck, Senior Member, IEEE

Abstract—This paper simulates the expected device performance and scaling perspectives of carbon nanotube (CNT) field-effect transistors with doped source and drain extensions. The simulations are based on the self-consistent solution of the three-dimensional Poisson–Schrödinger equation with open boundary conditions, within the nonequilibrium Green’s function formalism, where arbitrary gate geometry and device architecture can be considered. The investigation of short channel effects for different gate configurations and geometry parameters shows that double-gate devices offer quasi-ideal subthreshold slope and drain-induced barrier lowering without extremely thin gate dielectrics. Exploration of devices with parallel CNTs shows that ON currents per unit width can be significantly larger than the silicon counterpart, while high-frequency performance is very promising.

Index Terms—Ballistic transport, carbon nanotubes (CNTs), nonequilibrium Green’s function (NEGF), technology computer-aided design (CAD).

I. INTRODUCTION

CARBON nanotubes (CNTs) represent a promising alternative to conventional silicon technology [1] for future nanoelectronics at the end of the International Technology Roadmap for Semiconductors (ITRS) [2]. Since the first work on the topic by Iijima [3], significant improvements have been achieved from the point of view of both technology and physical modeling. In particular, Heinz et al. [4] have demonstrated Schottky barrier CNT-FETs, where the modulation of the current is mainly determined by the field-induced modulation of the nanotube band structure at the CNT ends. Such a working principle, however, strongly limits the device performance. The ambipolar behavior, the poor control of the channel, and the possible degradation of electrical properties like \( I_{on}/I_{off} \) ratio deeply affect such kind of devices, especially nanotubes with large diameters, as shown in [5]. To obtain acceptable drain-induced barrier lowering (DIBL), it has been demonstrated that the downscaling of device dimensions has to follow particular rules, like maintaining the ratio between the channel length and the oxide thickness larger than 18 [5].

To alleviate these problems, different solutions have been proposed to achieve channel modulation of the barrier. Javey et al. [6] and Nosho et al. [7] have respectively shown, for example, that ohmic contacts can be obtained by choosing Pb or Ca as the metal for the contacts. Inducing charge in the source and drain regions is another possible solution, reported, for example, in [8] and [9]. Channel modulation has also been obtained [10], [11] by definition of multiple gates able to fix independent potentials both in the reservoirs and in the channel as well as transparent Schottky barrier at the contacts.

In a scenario where many geometries are feasible, adequate physical models and simulation tools are necessary not only to provide explanations to experimental results but also to define device guidelines for the fabrication of CNT FETs, with performance benchmarked against their mainstream silicon counterpart.

Gate-all-around CNT FETs have been studied [12], [13], where the three-dimensional (3-D) Poisson equation has been reduced to two dimensions because of the cylindrical symmetry of the electrostatic potential, and transport has been computed through the Landauer formalism. Such a coaxial geometry has also been adopted in [14], where the Poisson equation has been coupled with the nonequilibrium Green’s function (NEGF) formalism using the uncoupled mode space approach, which enables the computation of transport for a small number of electron subbands with a small computational cost.

However, planar gate structures are more attractive because of their simpler fabrication technology, and all the experimental data discussed above are on planar geometries.

A full 3-D approach has been followed in [15], where the Poisson equation has been solved using the method of moments. Such a method provides the advantage of requiring the computation of the Poisson equation only in regions where charge is not zero, with the drawback that it is practically impossible to treat more than two different dielectric constants.

In this paper, we focus on realistic and experimentally relevant CNT-FETs with doped source and drain extensions,
and evaluate their performance against the requirements at the end of the ITRS. To this purpose, we have developed a code able to solve full band Schrödinger equations with open boundary conditions in the NEGF framework. Such a module has been included in our 3-D Poisson solver NANO.TCAD vIDES [16], which can deal with very general structures, since it does not take advantage of particular symmetries and can consider structures in which both CNTs and conventional semiconductors are simultaneously present.

Our realistic simulations show that CNT-FETs are very attractive for (1) their capabilities of suppressing short channel effects; (2) driving high ON currents per unit length; (3) providing $I_{on}/I_{off}$ ratio required by the ITRS for the 15-nm gate length, and (4) their potential for terahertz applications. On the other hand, we shall show that the subthreshold slope deteriorates for a gate voltage close to zero due to the filling of hole states in the channel.

II. SIMULATION APPROACH

The potential profile in the 3-D simulation domain obeys the Poisson equation

$$\nabla [\epsilon(\vec{r}) \nabla \phi(\vec{r})] = -q [p(\vec{r}) - n(\vec{r}) + N_D^{\uparrow}(\vec{r}) - N_A^{\downarrow}(\vec{r}) + \rho_{fix}]$$

(1)

where $\phi(\vec{r})$ is the electrostatic potential, $\epsilon(\vec{r})$ is the dielectric constant, $N_D^{\uparrow}$ and $N_A^{\downarrow}$ are the concentrations of ionized donors and acceptors, respectively, and $\rho_{fix}$ is the fixed charge. The electron and hole concentrations ($n$ and $p$, respectively) are computed by solving the Schrödinger equation with open boundary conditions by means of the NEGF formalism [17]. A tight-binding Hamiltonian with an atomistic ($p_z$ orbitals) real-space basis [18] has been used with a hopping parameter $t = 2.7$ eV.

Green’s function can then be expressed as

$$G(E) = [EI - \Sigma_S - \Sigma_D]^{-1}$$

(2)

where $E$ is the energy, $I$ the identity matrix, $H$ the Hamiltonian of the CNT, and $\Sigma_S$ and $\Sigma_D$ the self-energies of the source and drain, respectively. As can be seen, transport is assumed here to be completely ballistic.

The considered CNTs are all zigzag nanotubes, but the proposed approach can be easily generalized to nanotubes with a generic chirality since the required changes involve only the Hamiltonian matrix. Once the length and the chirality of the nanotube are defined, the coordinates in the 3-D domain of each carbon atom are computed [19]. After that, the 3-D domain is discretized so that a grid point is defined in correspondence with each atom, while a user-specified grid is defined in regions not including the CNT.

A point charge approximation is assumed, i.e., all the free charge around each carbon atom is spread with a uniform concentration in the elementary cell including the atom. Assuming that the chemical potential of the reservoirs is aligned at the equilibrium with the Fermi level of the CNT, and given that there are no fully confined states, the electron concentration is

$$n(\vec{r}) = 2 \int_{E_i}^{+\infty} dE \left[ |\psi_S(E, \vec{r})|^2 f(E - E_F) + |\psi_D(E, \vec{r})|^2 f(E - E_{F_D}) \right]$$

(3)

while the hole concentration is

$$p(\vec{r}) = 2 \int_{-\infty}^{E_i} dE \left[ |\psi_S(E, \vec{r})|^2 [1 - f(E - E_F)] + |\psi_D(E, \vec{r})|^2 [1 - f(E - E_{F_D})] \right]$$

(4)

where $\vec{r}$ is the coordinate of the carbon site, $f$ is the Fermi–Dirac occupation factor, $|\psi_S|^2(|\psi_D|^2)$ is the probability that states injected by the source (drain) reach the carbon site ($\vec{r}$), and $E_{F_S}$ ($E_{F_D}$) is the Fermi level of the source (drain). The current has been computed as

$$I = \frac{2e}{h} \int_{-\infty}^{+\infty} dE \left[ T(E) [f(E - E_{F_S}) - f(E - E_{F_D})] \right]$$

(5)

where $q$ is the electron charge, $h$ is Planck’s constant, and $T(E)$ is the transmission coefficient computed as [17]

$$T = -\text{Tr} \left[ \left( \Sigma_S - \Sigma_S^i \right) G \left( \Sigma_D - \Sigma_D^i \right) G^i \right]$$

(6)

where $\text{Tr}$ is the trace operator. We have to point out that with the present model, we only deal with the one-dimensional (1-D) transport between source and drain reservoirs, while the leakage gate current has not been taken into account. For the considered devices with channel length of a few nanometers, it can be shown that the gate current is negligible with respect to the drain current.

From a numerical point of view, Green’s function is computed by means of the recursive Green’s function (RGF) technique [20], [21]. Particular attention must be put in the definition of each self-energy matrix, which can be interpreted as a boundary condition of the Schrödinger equation. In particular, in our simulation, we have considered a self-energy for semi-infinite leads as boundary conditions, which enables to consider the CNT as connected to infinitely long CNTs at its ends.

We have to point out that the computation of the self-energy is quite demanding. To achieve faster results, we have followed the approach proposed in [22], which we have verified to be four times faster than a simple underrelaxation method.

We have to point out that, using a real-space basis, the computed current takes into account intraband and interband tunneling, since, as compared to the mode space approach, all the bands of the nanotube are considered simultaneously.

From a numerical point of view, the nonlinear system has been solved with the Newton–Raphson (NR) method with the Gummel iterative scheme. In particular, the Schrödinger equation is solved at the beginning of each NR cycle of the Poisson equation, and the charge density in the CNT is kept
constant until the NR cycle converges (i.e., the correction on the potential is smaller than a predetermined value). The algorithm is then repeated cyclically until the norm of the difference between the potential computed at the end of two subsequent NR cycles is smaller than a predetermined value.

Some convergence problems, however, may be encountered when using this iterative scheme. Indeed, since the electron density is independent of the potential within an NR cycle, the Jacobian is null for points of the domain including carbon atoms, losing control over the correction of the potential. We have used a suitable expression for the charge predictor to give an approximate expression for the Jacobian at each step of the NR cycle. To this purpose, we have used an exponential function for the predictor [22]. In particular, if \( n \) is the electron density as in (3), the electron density \( n_i \) at the \( i \)th step of the NR cycle can be expressed as

\[
    n_i = n \exp \left( \phi_i - \tilde{\phi} \right) / V_T
\]

where \( \tilde{\phi} \) and \( \phi_i \) are the electrostatic potentials computed at the first and \( i \)th steps of the NR cycle, respectively, and \( V_T \) is the thermal voltage. The same considerations follow for the hole concentration. Since the electron density \( n \) is extremely sensitive to small changes of the electrostatic potential between two NR cycles, the exponential function acts in the overall procedure as a damping factor for charge variations. In this way, convergence has been improved in the subthreshold regime and in the strong inversion regime. Convergence is still difficult in regions of the device where the charge in the nanotube is not compensated by a fixed charge, like in the case of bound states in the channel, where the right-hand term of the Poisson equation is considerably large.

III. RESULTS AND DISCUSSIONS

We first consider a (11, 0) CNT embedded in SiO\(_2\) with a diameter \( d \) of 0.9 nm, an undoped channel of varying length \( L \), and n-doped CNT extensions of 10 nm at the source and drain ends (Fig. 1). The CNT extensions have on average \( f \) donated electrons for each carbon atom. \( f \) is the product of the stoichiometric ratio of donors times the average number of electrons donated by each donor.

As a first attempt to study CNT-FET performance, we considered dc properties, evaluating the devices in terms of short channel effects and \( I_{on} \) and \( I_{off} \) currents. We first consider short channel effects for different gate layouts (single, double, and triple gates) for the same channel length \( L = 15 \) nm (Fig. 2).

Subthreshold swing \( S \) and DIBL as a function of oxide thickness for \( L = 15 \) nm and for different gate layouts. \( f = 10^{-3} \).

![Fig. 2. Transversal cross section of the CNT-FETs with different gate structures: (a) single gate, (b) DG, and (c) triple gate. Null Neumann boundary conditions are imposed on lateral ungated surfaces.](image)

![Fig. 3. (a) Subthreshold slope and (b) DIBL as a function of oxide thickness for \( L = 15 \) nm and for different gate layouts. \( f = 10^{-3} \).](image)
at 7 nm, while smaller channel lengths suffer from excessive degradation of the gate voltage control over the channel.

From here on, we will focus on a DG structure with $t_{ox} = 2$ nm and the cross section shown in Fig. 2(b). Fig. 5(a) shows the ON current $I_{on}$ per unit width as a function of channel length, assuming a lateral dielectric separating adjacent nanotubes of 2 nm. In Fig. 5(b), $I_{on}$ is plotted as a function of the nanotube diameter for a device with channel length equal to 7 nm.

Short channel effects become more important as the channel length is decreased; and at the same biasing conditions, shorter devices show larger $I_{on}$ currents, since lowering of the channel barrier occurs. Moreover, as far as the CNT diameter is increased, quantized states along each atom ring are closer in energy so that more subbands participate in electron transport, increasing the channel conductance.

Fig. 6(a) shows the ON current as a function of the normalized tube density per unit length $\rho = d/T$, where $T$ is the distance between the center of two nanotubes, as illustrated in the inset of Fig. 6(a). All the results show that CNT-FETs can drive significant currents. As compared with the ITRS requirements, in the case of the most densely packed array, the current per unit length is almost seven times larger than that expected for high-performance devices at the 32-nm technology node (hp32: effective gate length equal to 13 nm), and six times for the 22-nm technology node (hp22: effective gate length equal to 9 nm). In our opinion, such factors are attributable to the fully ballistic transport regime we have assumed and to the zero series resistance of the contacts. As a consequence, our results have to be considered as an upper limit for CNT performance.

The OFF current ($I_{off}$), defined as the current obtained for $V_{GS} = 0$ V and $V_{DS} = 0.8$ V, is 15 times larger than that required for both hp32 and hp22 nodes [Fig. 6(b)]. However, as shown in Fig. 7(a), where the OFF current is plotted as a function of the normalized tube density, as the density decreases, the current in the OFF state also decreases, so for a tube density smaller than $8 \times 10^{-2}$, we obtain a $I_{on}/I_{off}$ ratio larger than that required by the ITRS for the hp32 [Fig. 7(b)].

As also observed in [23], the degradation of the OFF current is due to the presence of bound states in the valence band, which,
for high doping and for large drain-to-source voltages, are occupied by holes tunneling from the drain reservoir [Fig. 8(a)]. For smaller $V_{DS}$, bound states are quite far from the drain Fermi level [Fig. 8(b)], so the linear behavior in the semilog plot of the transfer characteristics in the subthreshold regime is almost recovered, as shown in Fig. 9.

Another interesting effect due to bound states is that, in the negative gate voltage regime, when band-to-band tunneling occurs at both source and drain ends, the current increases for a negative gate voltage (Fig. 10). For a larger $V_{DS}$, this effect requires a larger negative $V_{GS}$ to be observed since the larger the drain-to-source voltage, the stronger the influence of the bound states in the valence band, which act against the activation of the band-to-band tunneling process. In addition, as can be seen in the case of $V_{DS} = 0.1$ V, for negative gate voltages, resonant states appear. Since in the considered CNTs the band gap is close to 1 eV, the above considerations suggest that such an effect could also limit the performance of silicon devices in the deca nanometer regime. Let us note that in such a band-to-band tunneling regime, the Coulomb blockade effects are relevant and not included in the NEGF approach. NEGF results can still give a qualitative understanding of CNT behavior for very negative gate voltages.

Fig. 11 shows the transconductance $g_m$ as a function of gate voltage for devices with channel lengths of 5, 7, and 10 nm computed for $V_{DS} = 0.8$ V. The transconductance peaks are

![Fig. 8. Density of states computed for the device with $L = 7$ nm, for $V_{GS} = 0$ V, and (a) $V_{DS} = 0.8$ V and (b) $V_{DS} = 0.5$ V, as a function of the energy and the coordinate along the nanotube axis. Dashed lines are in correspondence of the source and drain Fermi level.](image1)

![Fig. 9. Transfer characteristics for DG CNT-FET with $L = 7$ nm for $V_{DS} = 0.5$ V and $V_{DS} = 0.8$ V. $t_{ox} = 2$ nm, $f = 5 \times 10^{-3}$.](image2)

![Fig. 10. Transfer characteristics for DG CNT-FET with $L = 7$ nm for $V_{DS} = 0.5$ V and $V_{DS} = 0.1$ V. $t_{ox} = 1$ nm, $f = 5 \times 10^{-3}$.](image3)

![Fig. 11. Transconductance as a function of gate voltage for DG CNT-FETs with different channel lengths: $L = 5, 7,$ and $10$ nm; $t_{ox} = 2$ nm; $V_{DS} = 0.8$ V; and $f = 5 \times 10^{-3}$.](image4)
in correspondence with the gate voltage at which the first 1-D subband crosses the source Fermi level. As can be seen, a large transconductance is obtained for the all three considered devices.

We also considered the impact of the donated electron fraction $f$ on device performance. Fig. 12 shows the current for $V_{DS} = 0.5 \, \text{V}$ and $V_{GS} = 0 \, \text{V}$ as a function of $f$ for two DG CNT-FETs with $t_{ox} = 2 \, \text{nm}$ and $L = 7$ and $15 \, \text{nm}$. For such applied voltages, we verified that the interband tunneling currents are zero and the net current is due only to thermionic and intraband tunneling. As can be seen, for the $L = 7 \, \text{nm}$ device, the current is extremely sensitive to $f$, and a small variation in the number of ionized atoms in the source and drain extensions can determine variations of the current of almost two orders of magnitude. However, as the channel length is increased, such effect is weakened since the field generated by uncompensated donors in the reservoirs is less effective in lowering the channel barrier.

Since the number of donors is of the order of tens, current dispersion due to random dopant fluctuations can be problematic as shown in Fig. 12, where the current for a $L = 7 \, \text{nm}$ (7,0) CNT is shown. In this case, the number of atoms in the source and drain extensions has been decreased by a factor of 11/7, and the current in the OFF state varies in this case by almost three orders of magnitude for the same range of $f$.

We now focus our attention on the switching and high-frequency performance of CNTs. The typical figure of merit for digital applications is the intrinsic device speed, defined as $\tau = C_G V_{DD}/I_{on}$, where $V_{DD}$ is the supply voltage and $C_G$ is the differential gate capacitance for $V_{GS} = 0.8 \, \text{V}$ [Fig. 13(a)]. For our considered DG structure with common gate driving, if stray capacitances are negligible, $C_G$ can be computed as the derivative of the charge in the CNT with respect to the gate voltage. This quantity is typically used to estimate the time it takes an inverter to switch, when its output drives another inverter, represented as a load capacitance $C_L$, as shown in the inset of Fig. 13(a). Compared to the ITRS requirements for the hp22 technology node, the obtained $\tau$ is at least 12 times faster.

CNT-FETs also show good potential for terahertz applications [25], [26]. In Fig. 13(b), the cutoff frequency defined as $f_T = g_m/2\pi C_G$ is shown as a function of channel length: $f_T$ is of the order of tens of terahertz, and the values obtained by simulations are larger than those found in [25], where the gate capacitance is overestimated, since the electrostatic capacitance has been modeled as that of an infinite metallic cylinder over an infinite conducting plane. As a word of caution, we must consider that additional stray capacitances could reduce the estimated $f_T$.

IV. CONCLUSION

We have developed a novel 3-D NEGF-based simulation tool for arbitrary CNT-FET architectures, which has enabled us to investigate the performance perspectives of CNT-FETs from an engineering point of view.

We have, in principle, demonstrated that random distribution of dopants in the reservoirs can significantly affect device performance and degrade current in the OFF state by several orders of magnitude. We have shown that DG structures exhibit very small short channel effects even with a rather thick silicon oxide gate dielectric (5 nm), and still acceptable subthreshold swing and DIBL for devices with the channel length down to 7 nm.

The driving currents and transconductance are very good as compared to the ITRS requirements, while $I_{off}$ may pose some problems due to the presence of localized hole states in the channel. However, good $I_{on}/I_{off}$ can be achieved by reducing the tube density in the CNT-FET array, still satisfying ITRS requirements.

We have also shown that DG CNT-FETs are very promising for terahertz applications if stray capacitances can be maintained under control.

ACKNOWLEDGMENT

The authors would like to thank Prof. M. Lundstrom and Prof. M. Macucci for fruitful discussions and their support, and Prof. C. Fiegna for his suggestions on convergence improvement.
REFERENCES


Gianluca Iannaccone (M’98) received the M.S. and Ph.D. degrees in electrical engineering from the Università di Pisa, Pisa, Italy, in 1992 and 1996, respectively. In 1996, he took a permanent position as Researcher with the Italian National Research Council, and in the same year obtained a faculty position at the Electrical Engineering Department, Università di Pisa, first as an Assistant Professor, then, since January 2001, as an Associate Professor of electronics. He has coordinated many European and national projects involving multiple partners, and has acted as Principal Investigator in several research projects funded by public agencies at the European and national level and by private organizations. He has authored and edited over 100 papers published in journals with peer review and more than 60 papers in proceedings of international conferences. His interests include transport and noise modeling in nanoelectronic devices, development of technology computer-aided design tools, the investigation of limits to scaling down of CMOS technology, and the exploitation of nanoscale effects in alternative device structures. More recently, he is active in the design of extremely low power systems for radio frequency identification (RFID) and ambient intelligence scenarios.

Dr. Iannaccone is with the Technical committee of many international conferences and serves as Referee for leading journals in the fields of semiconductor physics, device electronics, and circuit design.

Gerhard Klimeck (S’94–M’96–SM’04) received the Ph.D. degree from Purdue University, West Lafayette, IN, in 1994 and the German electrical engineering degree from Ruhr-University Bochum, Bochum, Germany, in 1990. Previously, he was a Member of Technical Staff at the Central Research Lab, Texas Instruments Incorporated. He has been the lead on the development of NEMO 3-D, a tool that enables the simulation of tens-of-million atom quantum dot systems, and NEMO 1-D, the first nanoelectronic CAD tool. He was the Technical Group Supervisor for the Applied Cluster Computing Technologies Group and continues to hold his appointment as Principal Member at the Jet Propulsion Laboratory on a faculty part-time basis. Since December 2003, he has been the Technical Director of the Network for Computational Nanotechnology and a Professor of electrical and computer engineering at Purdue University, West Lafayette, IN. He leads the development and deployment of web-based simulation tools that are hosted on http://nanohub.org, a community website that is utilized by over 10,000 users annually. His work is documented in over 134 peer-reviewed publications and over 223 conference presentations. His research interest is in the modeling of nanoelectronic devices, parallel cluster computing, genetic algorithms, and parallel image processing.

Dr. Klimeck is a member of the American Physical Society,Eta Kappa Nu, and Tau Beta Pi.