Self-consistent Quantum Mechanical Treatment of the Ballistic Transport in 10 nm FinFET Devices Using CBR Method

Hasanur R. Khan, Denis Mamaluy, Dragica Vasileska

Department of Electrical Engineering
Ira A. Fulton School of Engineering, Arizona State University
Tempe, AZ 85287
hrkhan@asu.edu, mamaluy@asu.edu

ABSTRACT

As device size shrinks towards 10 nm feature size, ballistic transport and quantum interference effects are expected to play significant role in the operation of these nanoscale devices. It is, therefore, imperative to have a first principles device simulation tool that will take these effects into account. In this work, a fully quantum mechanical simulator based on Contact Block Reduction (CBR) method has been used to investigate the behavior of 10nm FinFET device in the ballistic regime of operation. Simulation results show the transformation of multiple channels into a single merged channel across the fin as the fin width is reduced gradually. Also we observe that short channel effects can be minimized by reducing the fin thickness which is evident from the device transfer characteristics for different fin thickness presented in this paper.

Keywords: FinFET, quantum transport, CBR method

1 INTRODUCTION

Scaling of conventional bulk-MOSFETs is approaching physical limits due to the upper bound imposed on the oxide thickness, S/D junction depth, etc. The double gate MOSFET has been proposed as a promising alternative structure [1] for future technology. With two gates controlling the entire fully depleted channel film, short channel effects (SCE) can be greatly suppressed. FinFET is attractive to researchers due to its quasi planar structure, better immunity to SCEs, automatic self-alignment of gates with each other and with S/D [2]. FinFET devices with gate length of 10 nm and oxide thickness of 1.7 nm have already been experimentally demonstrated [3].

At the same time, the quantum-mechanical transport simulation in 3D structures still presents a significant challenge due to profound computational difficulties. Although a large variety of computational methods developed over last decade to address this issue, many commonly used approaches are insufficient. For example, many methods cannot be used to accurately describe devices with more than two Ohmic contacts or have severe restrictions on device geometry (see *e.g.* Ref. in [4]). Another profound difficulty is that most presently available fully quantum mechanical transport methods suffer from extremely long simulation times and often can only be performed on large and expensive computer clusters.

In order to address these issues, a novel numerically efficient technique, based on the Green's function formalism, has been developed. The approach has been termed the Contact Block Reduction method [5,6] and recently has been extended to the self-consistent calculations. It allows one to calculate the transport properties of a two- or three-dimensional device that may have any shape, potential profile, and any number of leads.

In this work, we have applied the CBR method to investigate the ballistic quantum transport properties in nano-size FinFETs. Figure 1 depicts the geometry of FinFETs being simulated. It consists of vertical channels formed in opposite faces of Si fin controlled by self-aligned double gate [7,8]. The fin is usually intrinsic or lightly doped to avoid channel dopant fluctuation and threshold voltage sensitivity to fin width. The polysilicon n^+ gate has been assumed as a gate electrode. In this work gate length of 10 nm and oxide thickness of 1.75 nm have been used. The fin width is varied over a range from 12 nm down to 6 nm.

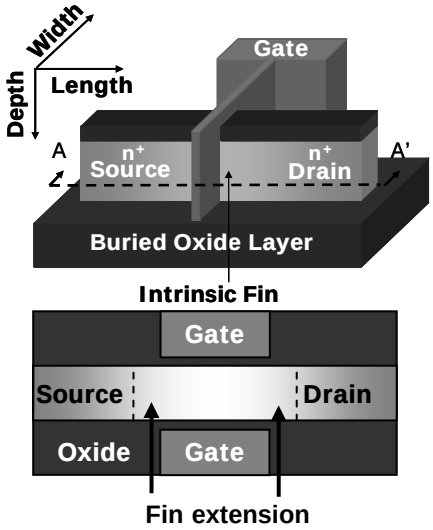

Figure 1.Top panel-3D view of FinFET device, Bottom panel-top view across A-A' cross section.

2 CBR METHOD

To calculate the transport properties of nano-size devices, one needs to take into account the fact that the Schrödinger equation for open systems does not, generally speaking, reduce to an "eigenvalue problem". Moreover, the standard band-structure calculation for a closed quantum system is simply insufficient to calculate any current at all: first of all the probability flow $q \sim \psi\nabla\psi^* - \psi^*\nabla\psi$ for any closed-system solution is zero (since the wave functions ψ of a closed system are real-valued), and secondly, it is obviously incorrect to attempt to calculate any current through a closed (i.e. non-interacting with the environment) system. There are several approaches available to take into account the "openness" of a realistic nano-device. Here we follow the Landauer-Büttiker description (see *e.g.* [9]), where the interaction of the system (i.e. device) with the environment is introduced via open boundary conditions. In particular, it is assumed that the external leads are reservoirs of carriers that are in equilibrium and can be described by a certain distribution function.

In the CBR method that is based on the Green's function formalism, quantities such as the transmission function and the charge density of the open system can be obtained from the eigenstates of a corresponding closed system $H^0|\alpha\rangle = E_\alpha|\alpha\rangle$ that needs to be calculated only once, and a solution of a very small linear algebraic system for every energy step. Importantly, it is shown in [6] that the calculation of relatively few eigenstates of the closed system is sufficient to obtain very accurate results. This makes it possible to apply the CBR method to complicated three-dimensional structures with arbitrary number of leads.

According to the Landauer-Büttiker formalism, the current $J_{\lambda\lambda'}$ from lead λ to lead λ' can be expressed in terms of the transmission function $T_{\lambda\lambda'}(E)$ and the distribution functions $f_\lambda(E)$ of the leads:

$$J_{\lambda\lambda'} = \frac{2e}{h}\int T_{\lambda\lambda'}(E)\big(f_\lambda(E) - f_{\lambda'}(E)\big)\,dE \qquad (1)$$

The transmission function can be obtained from the retarded Green's function $\mathbf{G}^R$ of the open device, and *the coupling to the leads* that is introduced via the self-energy matrix $\mathbf{\Sigma}$. The retarded Green's function of the open system $\mathbf{G}^R$ can be calculated (see *e.g.* [9]) using the Dyson equation and through the eigenstates $|\alpha\rangle$ of the closed system:

$$\begin{aligned}\mathbf{G}^R &= \mathbf{A}^{-1}\mathbf{G}^0,\ \mathbf{A} = \mathbf{1} - \mathbf{G}^0\mathbf{\Sigma},\\ \mathbf{G}^0 &= \big[E\mathbf{1} - \mathbf{H}^0\big]^{-1} = \sum_\alpha \frac{|\alpha\rangle\langle\alpha|}{E - E_\alpha}.\end{aligned} \qquad (2)$$

In work [5] it has been shown that the inversion of the matrix $\mathbf{A}$ can be easily performed using the property of the self-energy $\mathbf{\Sigma}$ in real space representation: it is non-zero only at boundary regions which are in ***contact*** with external leads. We denote these regions with index C, and the rest of the device with index D. After some algebraic manipulations [5] the Green's function matrix can be shown to have the following form:

$$\mathbf{G}^{\mathbf{R}} = \begin{bmatrix} \mathbf{A}_C^{-1}\mathbf{G}_C^0 & \mathbf{A}_C^{-1}\mathbf{G}_{CD}^0 \\ \mathbf{G}_{DC}^0\mathbf{\Sigma}_C\mathbf{A}_C^{-1}\mathbf{G}_C^0 + \mathbf{G}_{DC}^0 & \mathbf{G}_{DC}^0\mathbf{\Sigma}_C\mathbf{A}_C^{-1}\mathbf{G}_{CD}^0 + \mathbf{G}_D^0 \end{bmatrix} \qquad (3)$$

The ***small*** left-upper matrix block $\mathbf{G}_C^R = \mathbf{A}_C^{-1}\mathbf{G}_C^0$ fully determines the transmission function [5]. The particle density $n(\mathbf{r})$ can be obtained [6] using the following expression,

$$n(\mathbf{r}) = \sum_{\alpha,\beta}\langle\mathbf{r}\,|\,\alpha\rangle\langle\beta\,|\,\mathbf{r}\rangle\xi_{\alpha\beta}, \qquad (4)$$

where $\xi_{\alpha\beta}$ is the density matrix and is given by,

$$\begin{aligned}\xi_{\alpha\beta} &= \sum_{\lambda=1}^{L}\int \Xi_{\alpha\beta}^{(\lambda)}(E)\,f_\lambda(E)\,dE,\\ \Xi_{\alpha\beta}^{(\lambda)}(E) &= \frac{1}{2\pi}\frac{\mathrm{Tr}\Big(\big[|\beta\rangle\langle\alpha|\big]_C\,\mathbf{B}_C^{-1}\mathbf{\Gamma}_C^{(\lambda)}\mathbf{B}_C^{-1\dagger}\Big)}{\big(E - \varepsilon_\alpha + i\eta\big)\big(E - \varepsilon_\beta - i\eta\big)}\Bigg|_{\eta\to 0+}\end{aligned} \qquad (5)$$

Importantly, all the matrices in Eq. (5), including $\mathbf{B}_C = \mathbf{1} - \Sigma_C\mathbf{G}_C^0$ and $\mathbf{\Gamma}_C^{(\lambda)} = i\big(\mathbf{\Sigma}_C^{(\lambda)} - \mathbf{\Sigma}_C^{(\lambda)\dagger}\big)$, are of the size of the contact region only. Here, index (λ) denotes the contact region corresponding to lead λ only; otherwise it is assumed the entire contact region (i.e. all leads) should be taken into account.

Once the charge density $n(\mathbf{r})$ is known, the self-consistent solution for the open device should be obtained, that requires a repeated solution of the open-Schrödinger and Poisson equations. In this work the fully self-consistent calculation has been accomplished using the predictor-corrector approach [10], coupled to the CBR kernel, which allowed us to improve the convergence of a highly non-linear set of the coupled Schrödinger and Poisson equations. The details this self-consistent CBR scheme will be reported elsewhere.

3 SIMULATION RESULTS

In a FinFET, the fin width is one of the most important process variables as it determines the body thickness, which on the other hand, governs short-channel effects [11]. As the fin width increases gate control of the channel is worsened [12] which results in larger off-state leakage. For reasonable suppression of short channel effects fin thickness must be smaller than the gate length. For intrinsic fin, thickness of the fin needs to be even smaller.

In order to investigate these effects, we have simulated 10nm gate FinFET devices with different fin widths. Figures 2-4 depict the electron densities corresponding to the channel formation for different fin widths and were obtained for the drain bias of 0.1V and gate bias of 0.2V. In the case of fin width of 12 nm (Figure 2), the distinct channel forms vertically on each side wall of the fin along the gate, thus the device is in the double-channel mode. The center region of the fin remains almost intrinsic with very small electron density.

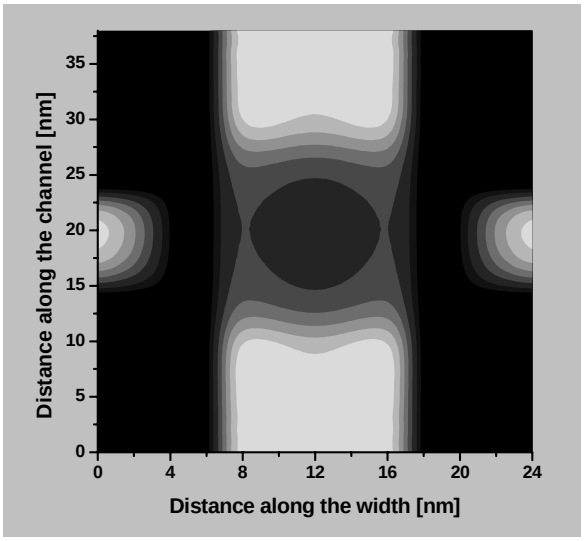

Figure 2: Electron density, fin width = 12 nm

As fin width decreases two channels gradually merge into a single channel across the fin. For a fin width of 10 nm two channels are almost merged except the center region of the fin (Figure 3).

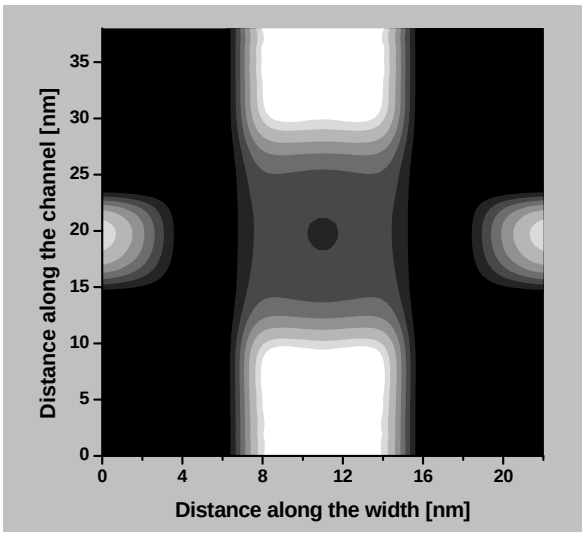

Figure 3: Electron density, fin width = 10 nm

Further, as fin width shrinks below 10 nm, the fin consists of a single channel resulting from volume inversion (Figure 4). For even smaller fin width, the channel profile does not change much, but suppression of short channel effects improves with a penalty of increased surface roughness. Also due to shrinking of fin width gate oxide has to be scaled down which on the other hand increases gate leakage as well as introduce fabrication issues.

The transfer characteristics for different fin thickness from 12 nm down to 6 nm are shown in Figure 5. The drain bias of 0.1 *V* has been used in the simulation.

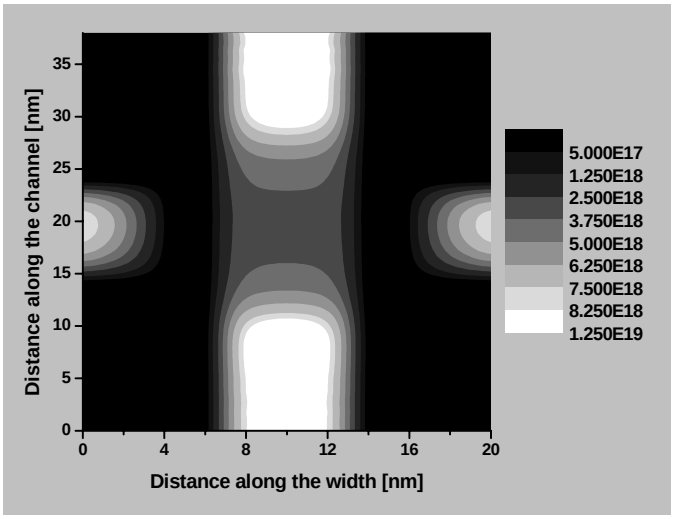

Figure 4: Electron density, fin width = 8 nm.
The color map for figures 2, 3 and 4 is shown here on the right side.

It is already evident from these simulation results that subthreshold characteristics can be significantly optimized by varying fin width. We note, however, that fin width can not be decreased arbitrarily to suppress short channel effects as discussed below.

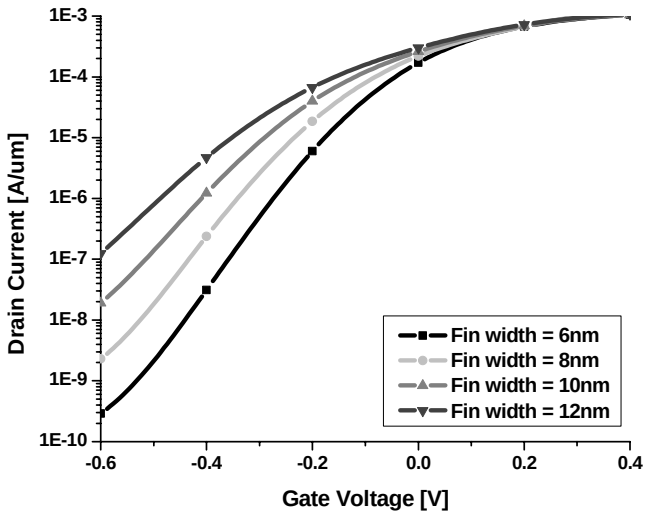

Figure 5: Transfer characteristics

The variation of subthreshold slope with fin width is shown in Figure 6. We find that as the fin width decreases, the subthreshold slope decreases almost linearly, which improves the controllability of short channel effects as expected. However, after some cutoff value (in this case it is around 6-7 nm), the subthreshold slope does not decrease anymore with decreasing fin width, but rather stays at a nearly constant value [13]. It is important to mention that while decreasing fin width improves the subthreshold behavior, on the other hand, the effect of surface roughness

starts to play significant role to degrade the device performance. A compromise must be drawn between these two factors to optimize the device behavior. Thus, for a more realistic FinFET simulation the effect of surface roughness should be included. That can be readily achieved in the CBR due to the real-space Green functions treatment. The effects of phonon scattering, while expected to be small in such ultra-scaled devices, also can be included into the simulator.

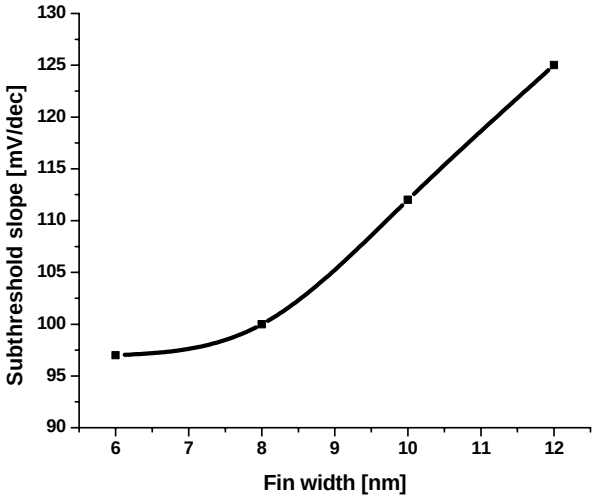

Figure 6: Subthreshold slope vs. fin width

4 CONCLUSION

FinFET is a novel device structure for future technology as it can overcome the scaling limit of conventional MOSFET towards 10 nm feature size. Short channel effects can be significantly reduced from the inherent structure of the device. With short gate length, the fin is usually much thinner and semi-classical simulation might not be sufficient to predict the device behavior. Fully quantum mechanical treatment becomes essential to investigate device behavior. In this work we developed fully quantum mechanical *self-consistent* CBR simulator to analyze characteristics of FinFET in subthreshold regime for different fin widths. The device turn-off behavior is examined by extracting sub-threshold slope for different fin widths. We find that while the subthreshold slope improves linearly with decreasing of the fin width, there is a limit beyond that any further shrinkage of the fin would not be beneficial, if the gate length is fixed. For example, for 10 nm gate length, the best turn-off performance can be achieved for about 6-8 nm fin width. To obtain a more realistic theoretical description of the device behavior beyond that limit, an approach including the surface roughness effects is necessary, which will be a subject of a forthcoming publication.

REFERENCES

[1] SIA Technology Roadmap for Semiconductors, (2003)

[2] H.S. Wong, K. Chan, and Y. Taur, IEDM Tech.Dig., 427, (1997)

[3] Bin Yu, Leland Chang, Shibly Ahmed et al., IEDM, pp. 251-254 (2002)

[4] D. K. Ferry and S. M. Goodnick, Transport in Nanostructures, (Cambridge University Press, Cambridge, UK, New York, (1997)

[5] D. Mamaluy, M. Sabathil, P. Vogl, J. Appl. Phys. **93**, 4628 (2003).

[6] D. Mamaluy, M. Sabathil, T. Zibold, D. Vasileska, P. Vogl, Phys. Rev. B **71**, 245321 (2005)

[7] D. Hisamoto, W-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, T-J. King, J. Bokor and C. Hu, IEEE Trans. Electron Dev., **47**, 2320, (2000)

[8] D. Hisamoto, W-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, T-J. King, J. Bokor and C. Hu, IEDM Tech. Dig., 1032, (1998)

[9] Supriyo Datta, Superlattices and Microstructures 28, 253 (2000).

[10] A. Trellakis, A.T. Galick, U. Ravaioli, J.H. Arends, Y. Saad, J. Appl. Phys. **81**, 3461 (1997)

[11] Chang, L., Choi, Y.-K., Ha, D., Ranade, P., Xiong, S., Bokor, J., Hu, C., and King, T.-J., *Proc. IEEE* **91** (11), 1860-1873,(2003)

[12] L. Chang, S. Tang, T.-J. King, J. Bokor, and C. Hu,

Int. Electron Devices Meeting Tech. Dig., pp. 719–722, (2000),

[13] Shiying Xiong and Jeffrey Bokor, IEEE Transactions on Electron Devices, vol.. 50, no. 11, November (2003).