Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Solid-State Electronics 49 (2005) 1510–1515
SOLID-STATE ELECTRONICS
www.elsevier.com/locate/sse

Quantum transport in ultra-scaled double-gate MOSFETs: A Wigner function-based Monte Carlo approach

V. Sverdlov [a,*], A. Gehring [b], H. Kosina [a], S. Selberherr [a]

[a] *Institute for Microelectronics, TU Vienna, Gusshausstr. 27–29, A-1040 Vienna, Austria*
[b] *AMD Saxony LLC and Co. KG, Wilschdorfer Landstrasse 101, 01109 Dresden, Germany*

Available online 19 September 2005

The review of this paper was arranged by Prof. F. Gamiz

Abstract

Source-to-drain current including tunneling in deca-nanometer double-gate MOSFETs is studied using a Monte Carlo approach for the Wigner transport equation. This approach allows the effect of scattering to be included. The subband structure is calculated by means of post-processing results from the device simulator MINIMOS-NT, and the contribution of the lowest subband is determined by the quantum transport simulation. Intersubband coupling elements are explicitly calculated and proven to be small in double-gate MOSFETs. The simulation results clearly show an increasing tunneling component of the drain current with decreasing gate length. For long gate length the semi-classical result is recovered.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Device simulation; Quantum transport; Wigner equation; Double-gate MOSFET

1. Introduction

Double-gate transistors are considered to be an attractive option to improve the performance of logic devices and overcome some of the difficulties encountered in further downscaling of bulk MOS field-effect transistors into the deca-nanometer regime [1]. When the channel length is reduced below approximately 25 nm, quantum effects such as direct source-to-drain tunneling through the barrier start affecting the device characteristics [2]. Frequently, ballistic transport is assumed which allows the device to be simulated using pure quantum-mechanical approaches [3–5]. However, with carrier mean free paths in the range of several nanometers [6], scattering-limited transport may still be dominant. A precise theory of ultra-scaled double-gate MOSFETs (DGMOSFETs) must therefore properly account for an interplay between the quantum nature of carrier propagation inside the channel and scattering processes. The effect of scattering has been assessed using classical Monte Carlo methods including a quantum correction [7] or the scattering rates of the two-dimensional electron gas [8]. The non-equilibrium Green's functions (NEGF) method [9] addresses the problem in the most consistent and complete way. However, due to its computational complexity, the method is frequently used in its reduced coherent version, which is equivalent to the solution of the Schrödinger equation with open boundary conditions. Introduction of scattering into the method requires the knowledge of the corresponding self-energies, which complicates computations significantly, allowing a successful solution only for some restricted class of model scattering mechanisms [10]. Inelastic processes are usually taken into account phenomenologically by introducing a "Büttiker probe" [11].

An alternative approach is based on the Wigner function formalism. It incorporates quantum-mechanical

* Corresponding author.
E-mail address: sverdlov@iue.tuwien.ac.at (V. Sverdlov).

0038-1101/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.sse.2005.07.013

effects in terms of a quantum scattering operator [12]. An advantage of the Wigner function approach is that it also includes all scattering mechanisms in a natural way via the Boltzmann scattering integrals, allowing a rigorous transport model to be developed, which accounting for both quantum interference phenomena and the scattering mechanisms.

It is essential that the Wigner equation formalism treats the scattering and quantum-mechanical effects on equal footing through the corresponding scattering integrals. The remaining collision-free propagation of the carriers is described by the Liouville operator acting on the Wigner function and is similar to that of the Boltzmann equation. It prompts for a solution of the Wigner equation with a Monte Carlo algorithm, by analogy to the solution of the Boltzmann equation. Such programs were recently realized in [12,13]. It was pointed out that, because of the kernel of the quantum scattering operator being not positively defined, the numerical weight of a particle trajectory increases rapidly in absolute value, and the numerical stability of a trajectory-based Monte Carlo algorithm becomes a critical issue. A multiple trajectories method was recently suggested [12] in order to overcome the difficulty. In this algorithm the problem of the scattering operator with the non-positive kernel and growing statistical weight of a single trajectory is rather addressed by creating a number of trajectories with positive and negative weights. Being formally equivalent to the former method, the later algorithm allows the annihilation of trajectories with similar statistical properties, which introduces a possibility to control the number of trajectories and as a result to obtain convergent results.

An additional convergence improvement may be achieved by an appropriate choice of a separation of the total potential energy into a classical part and a quantum-mechanical contribution. A substantial simulation speed-up was reported [14] while choosing a smooth classical component $V_c(x)$ as an output of a low-pass filter with the cut-off wave number $q_c \ll 2\pi/\Delta x$, where Δx is the grid step.

A common approximation used in simulations of DGMOSFETs is an adiabatic subband decomposition method, where it is believed that the intersubband coupling due to non-adiabatic effects should be small. A good agreement between the results of simulations found using the space-mode approach and the full-scale simulations [4] indicates that the subband decomposition may be a good approximation. However, recent studies have shown that even when the subband coupling is expected to be strong, results of current characteristic simulations using both of the foregoing approaches coincide within 10% [15]. This suggests that the coupling elements have to be evaluated in order to decide on their actual importance.

In this paper, we report on explicit evaluation of the subband coupling elements in DGMOSFETs, under non-equilibrium conditions when source–drain bias is applied. The potential in the structure is calculated self-consistently as an output of the MINIMOS-NT simulator [16]. It is shown that in ultra-scaled DGMOSFETs the coupling terms are few orders of magnitude smaller than the thermal energy and/or intersubband splitting, and the subband decomposition turns out to be an excellent approximation. Examples of Wigner function-based Monte Carlo simulations of the source–drain current including tunneling are presented in Section 3.

2. Effective Hamiltonian and transport model

The DGMOSFET is schematically shown in Fig. 1. An undoped $\langle 100 \rangle$ Si semi-conductor channel of thickness t is connected to the heavily doped source and drain contacts, with the dopant concentration set to 5×10^{19} cm^{-3}. The doping profile is chosen to be abrupt. Metal gates of length L with midgap work function are assumed. Gate electrodes are separated from the channel by thin layers of silicon dioxide, which in this study was chosen to be 1 nm in order to avoid appearance of an extra parameter. In the simulations the gate length L varies from 10 to 60 nm. The silicon film thickness t is in the range between 3 and 20 nm. The system is assumed to be infinite in z direction, such that the transversal motion in z direction is described by the conserved transversal momentum p_z and may be separated.

Due to the small silicon body thickness, the transversal carrier motion in y direction is quantized. The electron spectrum is split into well defined two-dimensional subbands. Each subband is characterized by the transversal wave function $\psi_n(y, x)$. We define the wave functions from the effective mass equation:

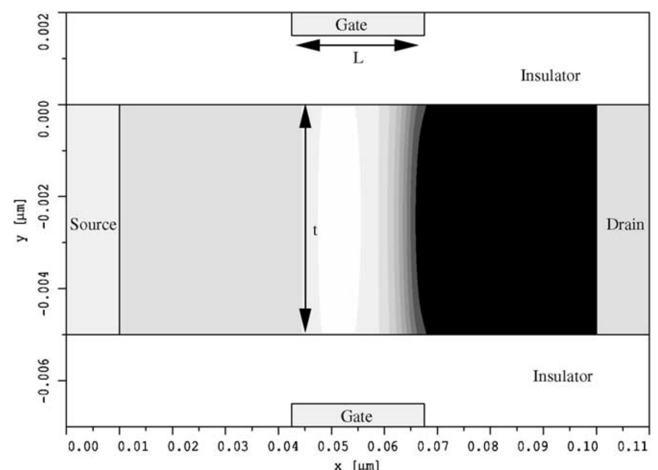

Fig. 1. Sketch of the 25 nm gate length double-gate MOSFET structure simulated by MINIMOS-NT. The contour plot shows the potential profile of the conduction band in the channel at the drain–source bias of 0.3 V and gate voltage 0.0 V.

$$\left[-\frac{\hbar^2}{2m_y}\frac{\partial^2}{\partial y^2}+U(y,x)\right]\psi_n(y,x)=\mathscr{E}_n(x)\psi_n(x,y), \tag{1}$$

where $U(y,x)$ is the self-consistent potential energy in the channel, and m_y is the effective mass in y direction. Note that due to the dependence of the longitudinal coordinate x still present in the potential profile $U(y,x)$, the eigenfunctions $\psi_n(y,x)$ and eigenenergies $\mathscr{E}_n(x)$ depend on x. Because of this dependence the subbands defined by $\psi_n(y,x)$ are not completely separable. However, in case of the DGMOSFET shown in Fig. 1, when mainly the silicon film boundaries determine the subband quantization, the subband coupling elements are expected to be small.

To rigorously evaluate the importance of subband coupling, we expand the full wave function $\Psi(x,y)$ using the complete set of subband functions $\psi_n(y,x)$ as

$$\Psi(x,y)=\sum_n\xi_n(x)\psi_n(y,x), \tag{2}$$

where coefficients $\xi_n(x)$ describe the longitudinal motion in transport direction within the nth subband.

It can be shown that the wave functions $\xi_n(x)$ satisfy the following system of coupled equations:

$$\left[-\frac{\hbar^2}{2m_x}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+(\mathscr{E}_n(x)+D_{nn}(x))\right]\xi_n(x)$$
$$=E\xi_n(x)-\sum_{m\neq n}\delta H_{nm}(x)\xi_m(x), \tag{3}$$

with open boundary conditions.

The intersubband coupling Hamiltonian is found to be

$$\delta H_{nm}(x)=S_{nm}(x)\frac{\partial}{\partial x}+D_{nm}(x), \tag{4}$$

where the coefficients $S_{nm}(x)$ and $D_{nm}(x)$ are defined as

$$S_{nm}(x)=-\frac{\hbar^2}{m_x}\int\mathrm{d}y\,\psi_n(y,x)\frac{\partial}{\partial x}\psi_m(y,x), \tag{5}$$

$$D_{nm}(x)=-\frac{\hbar^2}{2m_x}\int\mathrm{d}y\,\psi_n(y,x)\frac{\partial^2}{\partial x^2}\psi_m(y,x). \tag{6}$$

Because $\psi_n(y,x)$ are normalized, the coefficients satisfy the following relations:

$$S_{nm}(x)=-S_{mn}(x), \tag{7}$$

$$D_{nm}(x)=-D_{mn}+\frac{\hbar^2}{m_x}\int\mathrm{d}y\frac{\partial\psi_n(y,x)}{\partial x}\frac{\partial\psi_m(y,x)}{\partial x}. \tag{8}$$

It then follows that the diagonal terms D_{nn} are positively defined whereas $S_{nn}=0$ identically. With help of (7) and (8) one can also show that the matrix elements of the intersubband coupling Hamiltonian $\delta H_{nm}(x)$ possess the Hermitian property:

$$c_n^m\equiv\int\mathrm{d}x\,\xi_n^*(x)\delta H_{nm}(x)\xi_m(x)=(c_m^n)^*. \tag{9}$$

The potential profile inside the channel of DGMOSFETs calculated self-consistently for a drain–source voltage of 0.3 V and a gate bias of 0 V with help of MINIMOS-NT simulator is shown in Fig. 1. We have chosen this option of the potential calculation in the channel, in contrast to a self-consistent solution of the corresponding Poisson and Schrödinger equations, in order to account for strong dissipative processes due to electron–phonon scattering, which affect strongly the electron density and the potential profile in transport direction. To reproduce the carrier profile in the direction transversal to current correctly, we use MINIMOS-NT with a quantum correction potential introduced, which accounts for a proper electron density behavior at the channel boundaries. To be fully consistent with the Wigner Monte Carlo method, a self-consistent loop with the Wigner solver should be eventually implemented. This, however, imposes a major additional computational effort and goes beyond the scope of the current work.

The equipotential lines develop high curvatures at the drain end of the channel between the gates. It signifies a stronger dependence of the channel potential $U(y,x)$ on the longitudinal coordinate x and, as a consequence, a more pronounced dependence of subband wave functions (1) on x. The contour plot of the lowest subband function calculated for the same parameters as in Fig. 1 is shown in Fig. 2. As expected, the strong x dependence of the wave function is found in the same channel region where the channel potential $U(y,x)$ strongly changes with x.

The elements $D_{nm}(x)$ and $S_{nm}(x)$ of the coupling Hamiltonian are shown in Figs. 3–5, for several values of silicon thicknesses t and gate length L. As expected,

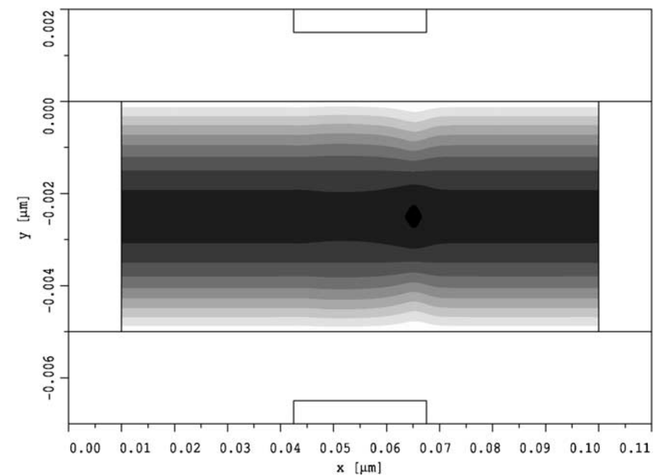

Fig. 2. Profile of the lowest subband wave function determined by (1). The parameters are the same as in Fig. 1.

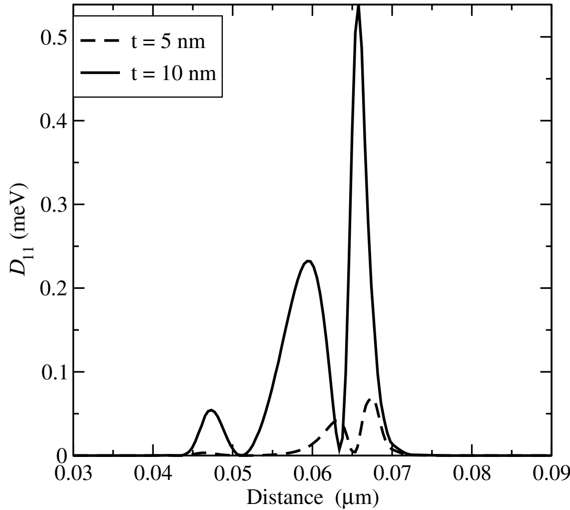

Fig. 3. Diagonal part $D_{11}(x)$ of the perturbation Hamiltonian calculated for gate length $L = 25$ nm, for two values of silicon body thickness. With the silicon thickness decreased, the diagonal part decreases rapidly.

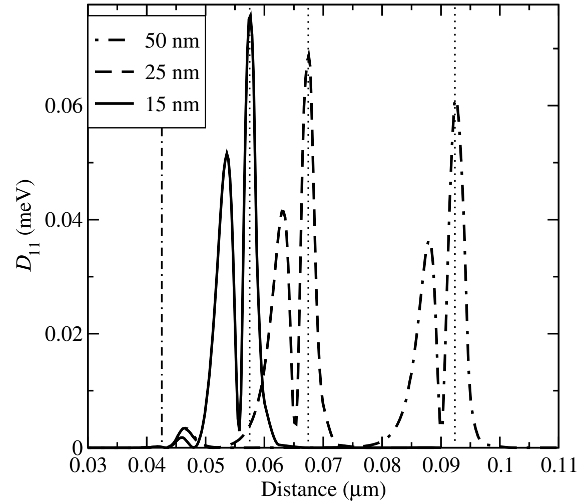

Fig. 5. Diagonal part of $D_{11}(x)$ of the perturbation Hamiltonian calculated for silicon body thickness $t = 5$ nm, for three different gate lengths. The dashed-dotted line shows the left edge and dotted lines the right edge of the gates, respectively.

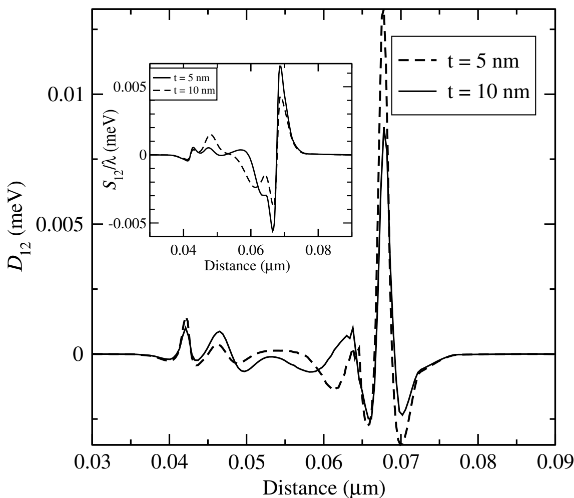

Fig. 4. Off-diagonal part $D_{12}(x)$ of the coupling Hamiltonian calculated for gate length $L = 25$ nm, for two values of the silicon body thickness t. The coupling increases slightly as the thickness is decreased. Inset: off-diagonal term $S_{12}(x)$, normalized to the thermal length $\lambda = \sqrt{2m_x k_B T}/\hbar$, for two values of body thickness.

all the coupling elements are substantially different from zero in the narrow channel segment where the channel potential and the transversal wave functions possess the strong longitudinal coordinate dependence.

Diagonal and off-diagonal coupling elements show quite a different behavior as a function of silicon body thickness t. The diagonal contribution $D_{nn}(x)$ is rapidly decreasing when the thickness of the silicon film shrinks (Fig. 3). Both contributions S_{nm} and D_{nm} to the off-diagonal part show an opposite tendency to grow slightly when the body thickness is reduced (Fig. 4 and inset). For fixed silicon film thickness, as the gate length L is decreased, the diagonal and off-diagonal contributions show similar dependence to increase. However, as one can see from Fig. 5, the L dependence of D_{nn} is quite weak, since the absolute increase is only about 15% when the channel length is reduced from 50 to 15 nm.

The first two subband profiles $\mathscr{E}_{Li}(x)$, $i = 1, 2$ with the transversal mass $m_y = 0.91m_0$, where m_0 the free electron mass, and the first subband profile $\mathscr{E}_{T1}(x)$ with $m_y = 0.19m_0$, are shown in Fig. 6. As can be seen from Fig. 4, the absolute value of the off-diagonal coupling elements $D_{nm}(x)$ and $S_{nm}(x)/\lambda$, where $\lambda = \sqrt{2m_x k_B T}/\hbar$ is the thermal length, is of the order of 10^{-2} meV and is extremely small compared to other characteristic

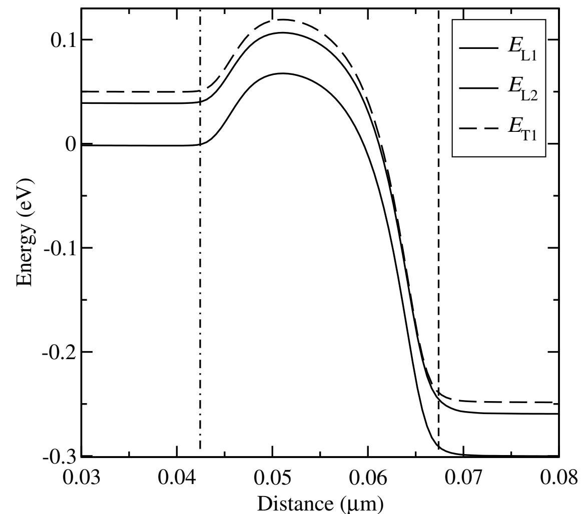

Fig. 6. Potential profile of three lowest subbands $\mathscr{E}_n(x)$ in $L = 25$ nm and $t = 5$ nm MOSFET. Only the lowest subband is substantially occupied with the carriers and thus participates in transport.

energies such as $k_B T$ or subband quantization energy (Fig. 6). It is thus justified to neglect the off-diagonal coupling $\delta H_{nm}(x)$ for the DGMOSFETs under consideration. If the coupling $\delta H_{nm}(x)$ is neglected, Eq. (3) coincides with the standard one-dimensional Schrödinger equation with the potential energy $V(x)$ equal to:

$$V(x) = \mathscr{E}_n(x) + D_{nn}(x). \tag{10}$$

Introducing the Wigner function $f_n(x,p,t)$ for each subband n and following the procedure outlined in [12], we obtain the one-dimensional Wigner equation for $f_n(x,p,t)$. The potential operator of the one-dimensional Wigner equation is given by the usual convolution integral [14], where the potential energy of electrons $V(x)$ is defined by (10).

Following [14], we introduce a spectral decomposition of the potential profile $V(x)$ into a slowly varying, classical component and a rapidly varying, quantum-mechanical component:

$$V(x) = V_{\mathrm{cl}}(x) + V_{\mathrm{qm}}(x). \tag{11}$$

This decomposition is conveniently carried out by applying a low-pass filter with a cut-off wave number $q_c \ll \pi/\Delta x$, where Δx is a grid step size. An example of such a decomposition is shown in Fig. 7. This separation of the total potential into classical and quantum-mechanical contribution improves significantly the Wigner Monte Carlo convergence. Electron–phonon scattering is taken into account through the conventional Boltzmann collision operator. Finally, for the DGMOSFETs under consideration, only the lowest subband was taken into account for Wigner transport calculations. This approximation is properly justified only for silicon thicknesses less than 5 nm. For thicker silicon bodies the higher subbands are getting increasingly populated so that the lowest subband approximation may provide only with qualitative understanding of the importance of coherent effects as compared to the classical simulations.

3. Results and discussion

The Wigner Monte Carlo simulations were carried for DGMOSFETs with gate lengths of 60, 25, 15, and 10 nm. The gate voltage was set to 1.0 V. The simulation domain was split into the quantum region which included the channel sandwiched between two classically treated parts in source and drain regions. The Wigner-type scattering due to the quantum potential leads to the generation of states with positive and negative statistical weights as explained elsewhere [12]. In order to enhance the efficiency of simulations in the quantum region and ultimately avoid situations when the convergence may not be achieved due to an insufficient number of negative weights necessary to treat correctly the Wigner function in the quantum region, trajectories with positive and negative weights were already generated in classical contacts with a predefined fixed ratio. The statistical compensation of positive and negative weights does not affect the expected value of estimators calculated in classical contacts and only leads to an increase of their variances. The trade-off for this algorithm is of course the CPU time, which grows proportionally to the number of trajectories used.

The profile of the two-dimensional carrier density along the channel and the typical output characteristics of DGMOSFETs are shown in Figs. 8 and 9, respectively. Results of the purely classical simulations when the Wigner-type quantum scattering was turned off are

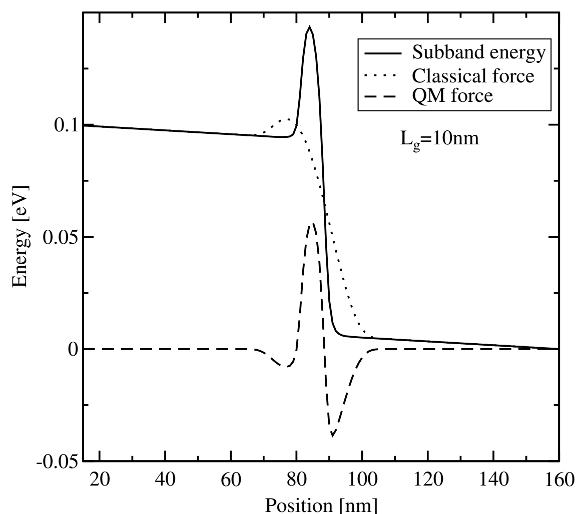

Fig. 7. Potential decomposition of the subband potential into a classical part $V_{\mathrm{cl}}(x)$ and a quantum-mechanical part $V_{\mathrm{qm}}(x)$ for the 10 nm gate length device.

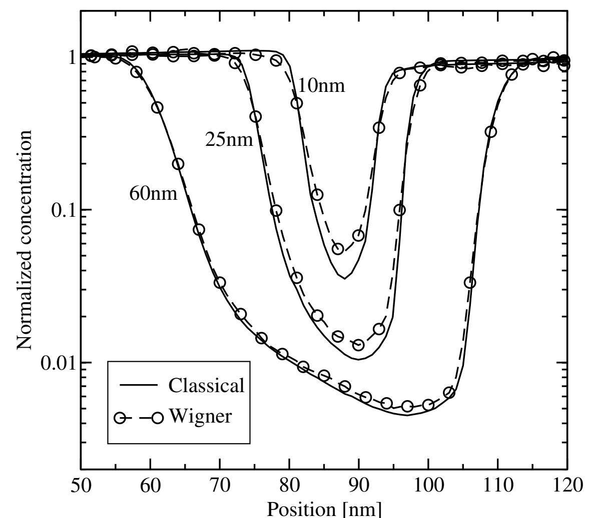

Fig. 8. The carrier concentrations in 10, 25 and 60 nm gate length device at 0.6 V drain bias for the classical and the Wigner Monte Carlo simulations.

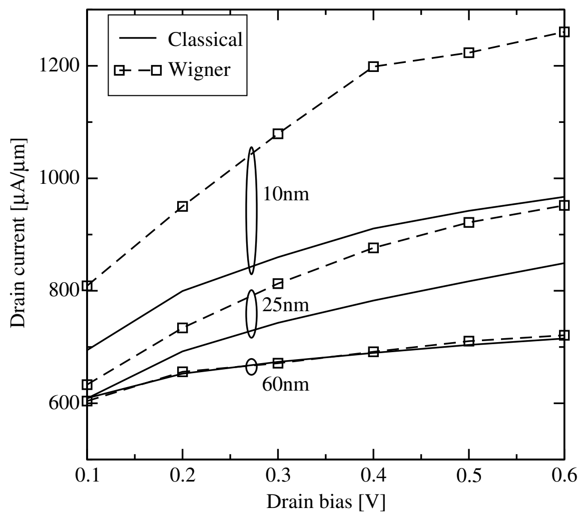

Fig. 9. Output characteristics of 10, 25, and 60 nm gate length device using classical and Wigner Monte Carlo. The additional source-to-drain tunneling current component is clearly visible for the short-channel devices.

also shown in both figures. For the gate length of 60 nm the classical density is identical to the one calculated with the quantum scattering included through the whole device. The difference starts to grow when the gate length decreased, with its maximum relative value acquired in the quantum region under the gate. For the gate length of 10 nm the actual density is already two times greater than the one found classically. This additional carrier concentration is of a quantum-mechanical nature and is created by the carriers, which are able to tunnel through the potential barrier created by the gates. The probability to find these carriers under the barrier in the channel is non-zero what leads to an additional quantum-mechanical contribution to the density.

Similar dependencies are observed for the drain-to-source current density in Fig. 9. For the long gate length the current found with help of the Wigner Monte Carlo method coincides with the thermo-ionic current computed classically within the statistical uncertainty represented by the symbol size. With the gate length decreased, the quantum current develops a significant increase as compared to the thermo-ionic classical contribution. The relative increase reaches already 50% for the gate length of 10 nm.

The current increase is also determined by its quantum-mechanical component, which is caused by the carriers being able to tunnel through the barrier. The potential barrier transmission coefficient grows rapidly with the gate length decreased, which explains the significant current increase over its classical contribution for the shorter gate length, in a complete analogy with the purely coherent case. The most important advantage of the Wigner Monte Carlo method over the latter one is that it allows to calculate the quantum-mechanical current and density contribution when dissipative processes caused by inelastic scattering are included.

4. Conclusions

The Wigner Monte Carlo method allows a direct comparison of semi-classical and quantum transport results. The method is used to account for the tunneling current in ultra-scaled double-gate MOSFETs. The subband decomposition method is applied. Intersubband coupling elements are explicitly computed and are shown to be small in DGMOSFETs. Devices of different gate lengths have been studied. Electron–phonon scattering processes are included via the corresponding scattering term. The Wigner calculations reproduce the semi-classical results for the long channel device and predict a significant source-to-drain tunneling component in the drain current at very short gate lengths.

Acknowledgments

This work has been partly supported by the Austrian Science Fund FWF, project P17285-N02, and the European Commission, project SINANO IST-506844.

References

[1] Colinge J-P. Solid-State Electron 2004;48:897.
[2] Wang J, Lundstrom M. Tech Dig IEDM 2002:715.
[3] Lundstrom M, Ren Z. IEEE Trans Electron Dev 2002;49:133.
[4] Venugopal R et al. J Appl Phys 2002;92:3730.
[5] Munteanu D, Autran J-L. Solid-State Electron 2003;47:1219.
[6] Jungemann C et al. Solid-State Electron 2004;48:1417.
[7] Khatawala GA, Winstead B, Ravaioli U. IEEE Trans Electron Dev 2003;50:2467.
[8] Gámiz F et al. Solid-State Electron 2004;48:937.
[9] Datta S. Electronic transport in mesoscopic systems. Cambridge University Press; 1995.
[10] Svizhenko A, Anantram MP. IEEE Trans Electron Dev 2003;50:1459.
[11] Büttiker M. Phys Rev Lett 1986;57:1761.
[12] Kosina H, Nedjalkov M, Selberherr S. J Comput Electron 2002;2:147.
[13] Shifren L, Ringhofer C, Ferry DK. IEEE Trans Electron Dev 2003;50:769.
[14] Gehring A, Kosina H. J Comput Electron 2005;4:67.
[15] Venugopal R et al. J Appl Phys 2004;95:292.
[16] Institut für Mikroelektronik, Technische Universität Wien, Austria, MINIMOS-NT 2.1 User's Guide, http://www.iue.tuwien.ac.at/software/minimos-nt, 2004.