Available online at www.sciencedirect.com
SCIENCE DIRECT®
Solid-State Electronics 48 (2004) 581–587
SOLID-STATE ELECTRONICS
www.elsevier.com/locate/sse

Modelling and simulation challenges for nanoscale MOSFETs in the ballistic limit

G. Curatola, G. Fiori, G. Iannaccone *

Dipartimento di Ingegneria dell'Informazione, Università degli studi di Pisa, Elettronica Info. e Telecomunicazioni, Via Diotisalvi 2, Pisa I-56122, Italy

Received 16 May 2003; received in revised form 9 July 2003; accepted 18 September 2003

Abstract

In this paper, we present the main issues and the modelling approaches for the simulation of nanoscale MOSFETs in which transport is dominated by ballistic electrons. We show that is indeed possible to compute in an accurate way the density of states in the channel in the case of quantum confinement without solving the complete two-dimensional Schrödinger equation. We are developing modelling tools that can be applied to several types of MOSFET structures: bulk, strained-Si and ultra-thin SOI MOSFETs, FINFETs, double gate MOSFETs and Schottky barrier MOSFETs. Here, results for silicon germanium and bulk silicon devices with channel length of 25 nm are presented. In the present form, tools are limited to the case of fully ballistic transport, which might be reached by the extremely scaled MOSFETs at end of the Roadmap.

1. Introduction

According to the 2002 update of ITRS [1,2], MOSFETs with effective channel length of 25 nm will enter large scale production in 2007, and MOSFETs with effective channel length of 9 nm in 2016, at the end of the Roadmap. Such devices are characterized by several aspects typical of the nanometer scale: short channel effects, quantum confinement in the channel, tunnel current through the gate dielectric, source-to-drain tunnel current, and far-from-equilibrium transport.

Already in present devices, a relevant fraction of electrons contributing to the drain current (in an *n*-MOSFET) traverse the channel without undergoing inelastic scattering. Such "ballistic" fraction will increase with scaling down, and it is predicted by some authors to be predominant over the fraction of electrons undergoing inelastic scattering in devices with channel length shorter than 30 nm [3].

Proper TCAD (Technology Computer Aided Design) tools for nanoMOSFETs should therefore consider "ballistic" electrons not as a perturbation of a quasi-equilibrium distribution, but, on the contrary, as mainstream electrons, and consider thermalized electron with a perturbative approach.

First analytical models of ballistic MOSFETs have been developed by Natori [4]. In addition, recent simulations performed with Monte Carlo semiclassical tools [5] or derived from a scattering theory of MOSFETs [6] exhibit significant differences with respect to simulations based on drift-diffusion or energy balance models.

Nanoscale MOSFETs also present a significant degree of quantum confinement in the channel, due to the high electric field in the vertical direction (perpendicular to the silicon–oxide interface) and to the very device structure (in the case of ultra-thin SOI or strained silicon).

A quantum simulation is therefore necessary, to take into account the separation of the density of states in well defined two-dimensional subbands, and degeneracy lift of silicon conduction band.

* Corresponding author. Tel.: +39-050-2217677; fax: +39-050-2217522.
E-mail address: g.iannaccone@iet.unipi.it (G. Iannaccone).

0038-1101/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.sse.2003.09.029

In recent years, MOSFETs realized with strained silicon channels on relaxed SiGe substrates have been proposed for improving high frequency and low power characteristics. Also in this case, the strained silicon layer has a thickness of the order of a few nanometers, and therefore causes significant quantum confinement of electrons in the channel. In addition, it is well known that electron mobility in strained silicon is significantly larger than in bulk silicon: it is therefore important to understand whether strained silicon provides comparable advantages also in the ballistic regime, where mobility has no physical meaning.

Other devices have been recently proposed as candidate structures for the 45 nm technology node and beyond, such as SOI devices, Schottky Barrier MOSFET, Double Gate MOSFETs, etc., which exhibit features that can only be addressed with a simulator capable of taking into account, in addition to ballistic transport, quantum confinement in the channel and multi-dimensional tunneling. From this point of view commercial TCAD programs are not adequate, neither concerning far from equilibrium transport, nor concerning quantum effects. On the other hand, in the scientific literature, several papers appeared that address in a different way parts of the problem:

- Through the introduction of a simple correction term (the "density gradient" approach) that takes into account quantum effects in an otherwise semiclassical drift-diffusion model [7–9].
- Through the solution of the 2D, or quasi-2D, Schrödinger equation in the channel, considering that transport in the channel can be described with a drift-diffusion model [10–12].
- Through the solution of 2D, or quasi-2D, of Schrödinger equation, with a fully ballistic transport model [13–17].
- Through the solution of a quasi-2D Schrödinger equation and of the quantum Liouville equation in each subband [18,19].

The method and the associated implementation presented in Refs. [18,19], are very promising, and apparently include all the relevant physics, except intersubband scattering. However, no data are presented as far as the computational efficiency is concerned. Other examples listed above are either poorly efficient from the computational point of view and mainly oriented at model validation, or address only devices with idealized (and simplified) structures or transport mechanism. However, the computational resources available on a low-end workstation would allow to implement a transport model capable to satisfy all modeling requirements indicated above, and to consider in a detailed way device architectures required to achieve the scaling objectives indicated in the Roadmap.

2. Model

In this paper we present an efficient code for the quantum simulation of ballistic MOSFETs. The code is based on the self-consistent solution in two dimensions of the Poisson equation, of the Schrödinger equation, and of the continuity equation for ballistic electrons. In addition, it takes properly into account the effect of strain on the band structure of silicon and on ballistic transport in the channel. The code is a remarkable result, since it demonstrates that quantum simulation of nanoMOSFETs, with realistic structure and doping profile, can be performed with a simple high-end PC, and does not require a huge processor cluster, as it has sometimes been proposed. The details of the model have already been described in Refs. [13,20] and will be only mentioned here.

Here, we want to discuss in some detail the choice of the occupation factor of electron states traveling from the source to the drain and viceversa, since a proper choice of such occupation factor, by itself, guarantees that the continuity equation for the ballistic current be obeyed (this is a great advantage, since there is no additional equation to be solved except Poisson–Schrödinger).

Let us consider the situation sketched in Fig. 1: the first subband profile in the longitudinal direction (from source to drain) has a peak close to the middle of the channel. This peak represents the energy of the barrier controlling the drain current: the maximum value of the subband (let us call it E_{sbmax}) is modulated via the gate voltage. If we assume that electrons do not lose energy, then we can assign to each electron state the occupation factor corresponding to the originating contact. Let us call $f(y,E)$ the occupation factor at coordinate y (in the longitudinal direction) and for energy E. If $E < E_{\mathrm{sbmax}}$ then $f(y,E) = f_{\mathrm{FD}}(E_{\mathrm{FS}},E)$ if y is on the left of the peak,

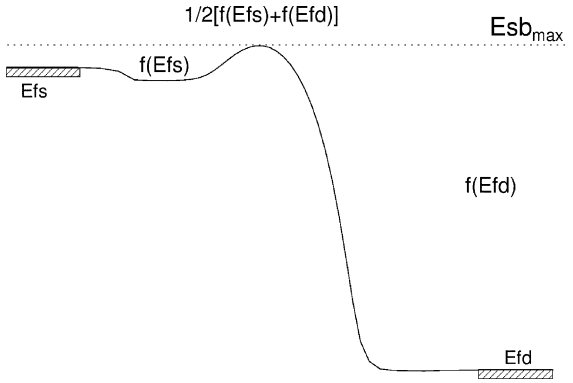

Fig. 1. Sketch of the subband profile along the channel, and of the occupation factor assigned to electron states as a function of energy and position.

$f(y,E) = f_{FD}(E_{FD},E)$ if y is on the right of the peak, where f_{FD} is the Fermi–Dirac occupation factor, E_{FS} and E_{FD} the Fermi energy of the source and drain reservoirs, respectively. Instead, if $E > E_{sbmax}$ then for each y and E there is a state traveling towards right (originating from the source) and a state traveling towards left (originating from the drain), therefore, the occupation factor is $f(y,E) = [f_{FD}(E_{FS},E) + f_{FD}(E_{FD},E)]/2$.

Since confinement is particularly strong in the vertical direction, we can decouple the two-dimensional Schrödinger equation in two one-dimensional equations in the vertical (x) and longitudinal (y) directions. In particular, the Schrödinger equation reads:

$$\left(-\frac{\hbar^2}{2m_x}\frac{\partial^2}{\partial x^2} - \frac{\hbar^2}{2m_y}\frac{\partial^2}{\partial y^2} \right) \Psi(x,y) + E_C(x,y)\Psi(x,y) = E\Psi(x,y), \quad (1)$$

where m_x and m_y are the effective masses along the x and y directions, respectively, E_C is the conduction band, E and Ψ are the energy and the wave function, respectively. In general, we can write $\Psi(x,y) = \psi(x,y)\chi(y)$, where ψ is the solution of the one-dimensional Schrödinger equation in the vertical direction in position y:

$$-\frac{\hbar^2}{2m_x}\frac{\partial^2}{\partial x^2}\psi(x,y) + E_C(x,y)\psi(x,y) = \widetilde{E}(y)\psi(x,y), \quad (2)$$

where $\widetilde{E}(y)$ is the energy eigenvalue. Since confinement is weak in the longitudinal direction, we approximate χ with a plane wave

$$\chi = \exp\left(\pm\frac{iy}{\hbar}\sqrt{2m_y[E - \widetilde{E}(y)]} \right). \quad (3)$$

In order to evaluate the error introduced by our approximations, we can substitute ψ and χ obtained from (2) and (3) in (1) and compute the parameter δ defined as

$$\delta = \max_y \left| \frac{-\frac{\hbar^2}{2m_x}(\psi\chi) + \psi\frac{\hbar^2}{2m_x}\chi}{[E - \widetilde{E}_1(y)]\psi\chi} \right| \quad (4)$$

δ represents the ratio of the residue of the Schrödinger equation to the difference between the total energy and the difference between the total energy and the first energy eigenvalue in the vertical direction, i.e., is a measure of the relative error introduced by our approximation. As can be seen in Fig. 2, for the so-called well tempered MOSFET with channel length 25 nm [21], δ is always below 10^{-4} in the case of continuous doping profile, and 10^{-3} in the case of discrete distribution of impurities, where band profiles are very rough.

If we consider only the first subband $E_{sb}(y)$, i.e. the lowest eigenvalues associated to the eigenstates $\Psi_l(x,y)$ obtained by solving the Schrödinger equation in the vertical direction with the longitudinal effective mass, the electron density n reads:

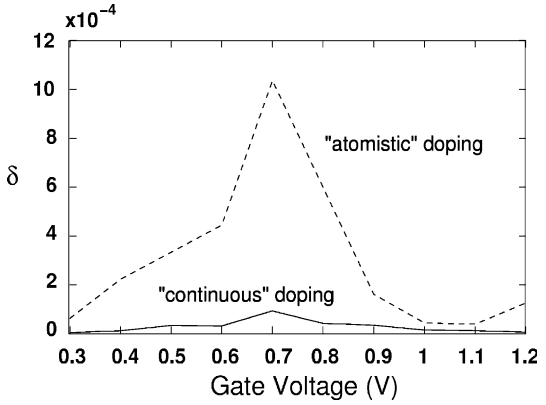

Fig. 2. Plot of the parameter δ defined in Eq. (4) as a function of the gate voltage for the 25 nm well tempered MOSFET defined in Ref. [21], in the case of continuous doping profile (solid line) or "atomistic" doping profile.

$$\begin{aligned} n(x,y) &= |\Psi_1(x,y)|^2 \int_0^{E_{sbmax}-E_{sb}(y)} D_1(E_y) f_{FD}(E_{FS/D}, E_y + E_{sb}(y))\,dE_y \\ &+ |\Psi_1(x,y)|^2 \int_{E_{sbmax}-E_{sb}(y)}^{\infty} D_1(E_y) \\ &\times \left[\frac{f(E_{FS}, E_y + E_{sb}(y)) + f(E_{FD}, E_y + E_{sb}(y))}{2} \right] dE_y \end{aligned} \quad (5)$$

where $D_l(E_y)$ is the longitudinal density of states and $E_{FS/D} = E_{FS}$ (E_{FD}) if y is on the left (right) of the peak.

The ballistic current density in a subband, per unit width, is given by the formula

$$\begin{aligned} J_1 = q\Gamma(1/2)\frac{\sqrt{2m_t kT}}{h^2} &\times \int_{E_{sbmax}-E_{sb}}^{\infty} \left[F_{-1/2}\left(-\frac{E_{sb} + E_y - E_{FS}}{kT} \right) \right. \\ &\left. - F_{-1/2}\left(-\frac{E_{sb} + E_y - E_{FD}}{kT} \right) \right] dE_y \end{aligned} \quad (6)$$

where m_t is the transversal mass, Γ is the Gamma function and $F_{-1/2}$ is the Fermi integral of order $-1/2$. The total current is obtained by summing up the above expression over all subbands.

In addition, our code takes properly into account the effect of strain on the band structure of silicon and on ballistic transport in the channel, using a model described in Ref. [22].

Here, we show how the code can be applied to the simulation of three different types of devices:

- A bulk-Si MOSFET with channel length of 25 nm and super-halo p-doping, proposed by Antoniadis as a benchmark device [21].
- A strained-Si MOSFET with the same doping profile proposed in [21], fabricated on the SiGe heterostructure shown in Fig. 3 (on the left). The strained silicon layer is 10 nm.
- A strained-Si MOSFET with an epitaxial SiGe p-doped layer instead of the implanted p-doping on the layer structure shown in Fig. 3 (right). The strained silicon layer is 10 nm, the p-doped epitaxial SiGe layer is 40 nm, with $N_A = 1.7 \times 10^{19}$ cm^{-3}, corresponding to the same impurity dose as the above mentioned devices.

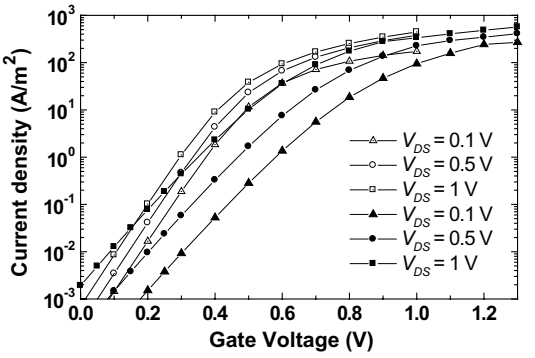

Fig. 4. Transfer characteristics of the bulk-Si 25 nm MOSFET computed with the quantum ballistic code (black symbols) and with Medici (white symbols, from Ref. [21]).

3. Results and discussion

The transfer characteristics of the bulk-Si 25 nm MOSFET computed with the ballistic model are plotted in Fig. 4, and are compared with results obtained from the commercial simulator MEDICI, in which quantum corrections are included (data from Ref. [21]). Complete solution of the Schrödinger equation, as can be seen, leads to a significant increase of the threshold voltage, while ballistic transport implies a significantly larger current in the on state, as well as larger threshold voltage since quantum confinement is fully taken into account. The subthreshold voltage swing appears to be larger in the case of ballistic transport. In Fig. 5, we show the computed electron density in the same device in saturation. Electrons in the channel are mostly in the first subband, which contributes to more than 95% of the total current in a broad range of operating voltages.

In strained-Si channels subband selection is even stronger: in Fig. 6 we compare the conduction band profile at inversion of the bulk Si and the epilayer-doped strained-Si MOSFET for $V_{GS} = 1$ V. The subbands corresponding to the minima along the y, z directions

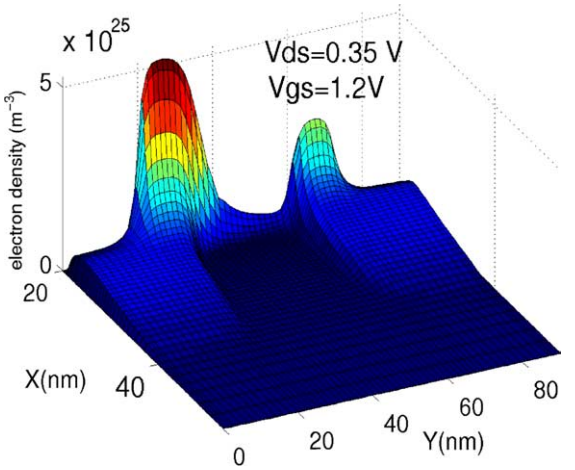

Fig. 5. Theoretical electron density in the bulk-Si 25 nm MOSFET for $V_{DS} = 0.35$ V, $V_{GS} = 1.2$ V.

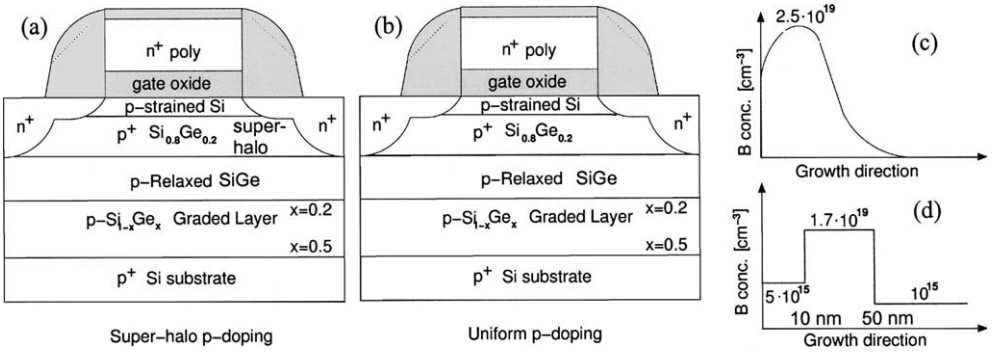

Fig. 3. SiGe MOSFETs considered (not in scale): Strained-Si MOSFET (a) with super-halo p-doping (c); Strained-Si MOSFET (b) with epitaxial SiGe p-doped layer (d).

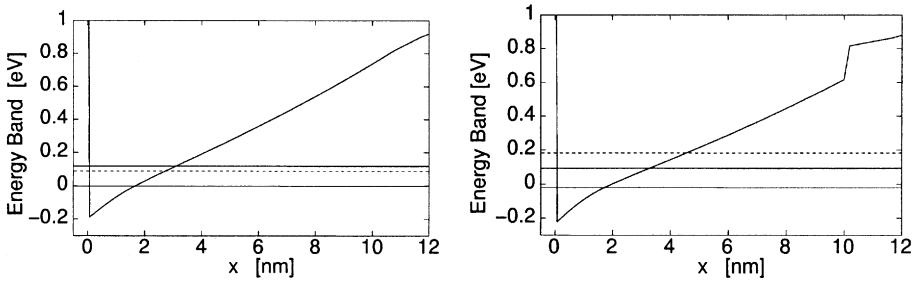

Fig. 6. Conduction band profile and subband energies the middle of the channel for the bulk-Si MOSFET (left) and for the strained-Si MOSFET with epi doping (right).

are significantly shifted towards higher energies and therefore are hardly populated in normal operating conditions.

The transfer characteristics and the transconductance of the bulk silicon MOSFETs, of the SiGe MOSFET with super-halo doping, and of the SiGe MOSFET with a doped epitaxial SiGe layer are compared in Fig. 7. The presence of the silicon–germanium heterostructure significantly affects the threshold voltage, which must be adjusted by properly tuning the doping profile and dose, but the current drive capability of the device seems not to be affected by the fact that the channel forms in strained silicon (curves in Fig. 7 seem only to be laterally shifted). This aspect needs to be investigated in deeper detail, since seems to suggest that the known mobility increase in strained silicon channels does not provide significant advantage in the ballistic regime. Actually subband selection, which is mainly responsible for the increased current drive on strained silicon, seems to be obtained also in conventional silicon channels in aggressively scaled down devices, when confinement due to the electric field is sufficiently strong. Complete output characteristics of the strained-Si MOSFETs are shown in Fig. 8, while results for the bulk silicon MOSFET can be found in Ref. [13].

The simulations shown suggest that the code is a versatile tool for addressing several different types of devices and materials. It allows us to analyze device performance in a transport regime in which the concept of mobility is not meaningful. It also allows to tune the threshold voltage and the transconductance by carefully adjusting the layer structure and the doping profiles. Further, extensive simulations and comparison with experimental results are needed to fully validate the code.

In order to address the needs of future technology nodes, the model has to be further improved. In particular, we are including the following features:

- Efficient multidimensional tunneling.
- Inelastic carrier scattering in the channel.
- Time dependent simulations for extracting the parameters important for compact model for RF design.

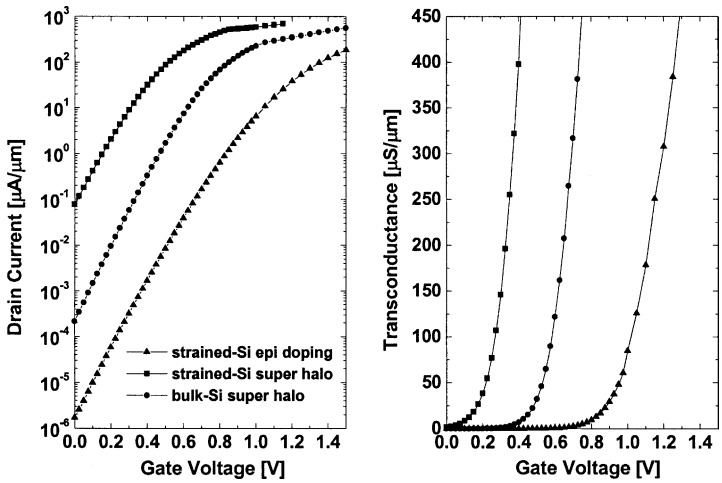

Fig. 7. Drain current and transconductance as a function of the gate voltage for $V_{DS} = 0.5$ V for the bulk-Si MOSFET (circles), the strained Si epilayer MOSFET (triangles) and the strained Si super-halo MOSFET (squares).

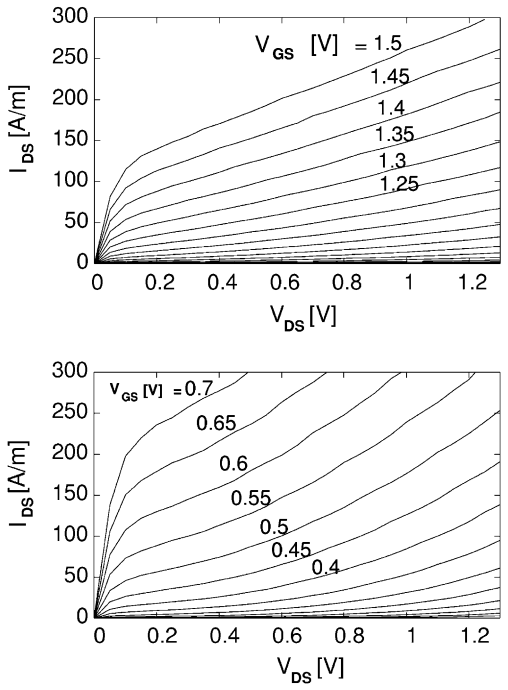

Fig. 8. Output characteristics of the strained-Si MOSFET with super-halo doping (above) and with doping of the epitaxial layer (below).

A version of the code extended to include tunnelling in each subband is already publicly available, and usable via a web interface [23]. For very short (<10 nm) channel lengths, instead, the computation of the two-dimensional transmission matrix is required. This feature can be straightforwardly included in the code, without requiring much larger computational resources.

The inclusion of a model for inelastic scattering in the channel, that keeps the simulation time under control is the real challenge of the project, and is still an open problem. The model presented in Refs. [18,19] is very interesting, but no information is available on the simulation time. One viable approach, according to us, is to treat inelastic scattering as a first-order perturbation to ballistic transport, completely reversing the traditional approach of models derived from drift-diffusion. This approach would also have the additional advantage of becoming more and more accurate, as device size is scaled down.

Acknowledgements

Support from the NANOTCAD project (EU Contract IST-1999-10828) and from the Fondazione Cassa di Risparmio di Pisa is gratefully acknowledged.

References

[1] International Technology Roadmap for Semiconductors 2001, Semiconductor Industry Association, San José, USA, Available from <http://public.itrs.net>.
[2] 2002 Update of the International Technology Roadmap for Semiconductors, Semiconductor Industry Association, San José, USA, Available from <http://public.itrs.net>.
[3] Frank DJ, Laux SE, Fischetti MV. Monte Carlo simulation of a 30 nm dual-gate MOSFET: How short can Si go? IEDM Tech Dig 1992:553–6.
[4] Natori K. Ballistic metal–oxide–semiconductor field-effect-transistor. J Appl Phys 1994;72:4879–90.
[5] Bude JD. MOSFET modeling into the ballistic regime, Proceedings of the International Conference on the Simulation of Semiconductor Processes and Devices (SISPAD) 2000. p. 23–6.
[6] Lundstrom M, Ren Z, Datta S. Proceedings of the International Conference of the Simulation of Semiconductor Processes and Devices (SISPAD) 2000. p. 1–5.
[7] Ancona MG et al. Density-gradient analysis of MOS tunneling. IEEE Trans Electron Dev 2000;47:2310–9.
[8] Wettstein A, Schenk A, Fichtner W. Quantum device simulation with the density-gradient model on unstructured grids. IEEE Trans Electron Dev 2001;48:279–84.
[9] Connelly D, Yu Z, Yergeau D. Macroscopic simulation of quantum mechanical effects in 2D MOS devices via the density gradient method. IEEE Trans Electron Dev 2002; 49:619–26.
[10] Spinelli A, Pacelli A, Lacaita A. Self-consistent 2D model for quantum effects in nMOS transistors. IEEE Trans Electron Dev 1998;45:1342–9.
[11] Abramo A, Cardin A, Selmi L, Sangiorgi E. Two-dimensional quantum mechanical simulation of charge distribution in silicon MOSFETs. IEEE Trans Electron Dev 2000; 47:1858–63.
[12] Pirovano A, Lacaita AL, Spinelli A. Two-dimensional quantum effects in nanoscale MOSFETs. IEEE Trans Electron Dev 2002;47:25–31.
[13] Fiori G, Iannaccone G. Modeling of ballistic nanoscale metal-oxide-semiconductor field effect transistor. Appl Phys Lett 2002;81:3672–4.
[14] Datta S. Nanoscale device modeling: The Green's function method. Superlattices Microstruct 2000;28:253–78.
[15] Ren Z, Venugopal R, Datta S, Lundstrom M, Jovanovic D, Fossum J. The ballistic nanotransistor: A simulation study. IEDM Tech Dig 2000:715–8.
[16] Ren Z, Venugopal R, Datta S, Lundstrom M. Examination of design and manufacturing issues in a 10 nm double gate MOSFET using nonequilibrium Green's function formalism. IEDM Tech Dig 2001:107–10.
[17] Svizhenko A, Anantram MP, Govindan TR, Biegel B, Venugopal R. Two-dimensional quantum mechanical modeling of nanotransistors. J Appl Phys 2000;91:2343–54.
[18] Balaban SN, Pokatilov EP, Fomin VM, Gladilin VN, Devreese JT, Magnus W, et al. Quantum transport in the cylindrical sub-0.1 μm silicon-based MOSFET. Solid-State Electron 2002;46:435–44.

[19] Croitoru MD, Gladilin VN, Fomin VM, Devreese JT, Magnus W, Schoenmaker W, et al. Quantum transport in a nanosize silicon-on-insulator metal–oxide–semiconductor field-effect transistor. J Appl Phys 2003;93:1230–40.

[20] Fiori G, Iannaccone G. The effect of quantum confinement and discrete dopants in nanoscale 50 nm *n*-MOSFETs: A three-dimensional simulation. Nanotechnology 2002;13(3): 294–8.

[21] Home page of the well tempered MOSFET at MIT: Available from <www.ece.mit.edu/Well>.

[22] Curatola G, Iannaccone G. Quantum confinement in silicon–germanium electron waveguides. Nanotechnology 2002;13:267–73.

[23] Iannaccone G, Curatola G. NANOTCAD2D on the PHANTOMS nanotechnology hub: Available from <www.phantomshub.com>.