Available online at www.sciencedirect.com

ScienceDirect

Microelectronics Reliability 47 (2007) 11–19

MICROELECTRONICS RELIABILITY

www.elsevier.com/locate/microrel

Introductory Invited Paper

Modeling current transport in ultra-scaled field-effect transistors ☆

V. Sverdlov, H. Kosina, S. Selberherr *

Institute for Microelectronics, TU Vienna, Gusshausstr. 27–29, A-1040 Vienna, Austria

Received 28 February 2006; received in revised form 1 March 2006
Available online 15 May 2006

Abstract

An overview of models used for the simulation of current transport in nanoelectronic devices within the framework of TCAD applications is presented. Modern enhancements of semiclassical transport models based on microscopic theories as well as quantum-mechanical methods used to describe coherent and dissipative quantum transport are specifically addressed. This comprises the incorporation of quantum corrections and tunneling models up to dedicated quantum-mechanical simulators, and mixed approaches which are capable of accounting for both, quantum interference and scattering. Specific TCAD requirements are discussed from an engineer's perspective and an outlook on future research directions is given.
© 2006 Elsevier Ltd. All rights reserved.

Contents

1. Introduction . . . 11
2. Semiclassical transport . . . 12
3. Quantum-ballistic transport . . . 13
4. Dissipative quantum transport . . . 15
5. Conclusions . . . 18
Acknowledgments . . . 18
References . . . 18

1. Introduction

The breathtaking increase in computational power and speed of integrated circuits in the past decades has been supported by the aggressive size reduction of semiconductor devices. This trend is expected to continue in the coming decade as predicted and institutionalized by the International Technology Roadmap for Semiconductors [1]. Today, when the 90 nm technology node with physical transistor gate lengths in the range of 40 nm is in mass production, the challenge is to introduce the 65 nm technology node already in a year. A new technology node is introduced every 3 years, with a long-term projection of the 22 nm node to be in mass production by the year 2016. A possibility to build metal-on-insulator field-effect transistors (MOSFETs) with even shorter gate lengths has been successfully established after the 6 nm gate length transistor has been demonstrated in research labs [2,3]. From a theoretical viewpoint even a few nm gate length device

☆ An earlier version of this paper was published in the Proceedings of the 2005 International Conference on Electron Devices and Solid-State Circuits (EDSSC 2005), Hong Kong, 19–21 December 2005, pp. 385–390.

* Corresponding author. Fax: +43 1 58801 36099.
E-mail addresses: sverdlov@iue.tuwien.ac.at (V. Sverdlov), Selberherr@TUWien.ac.at (S. Selberherr).

0026-2714/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.microrel.2006.03.009

has been predicted to be functional [4,5]. Nevertheless, emerging outstanding technological challenges related to different aspects of MOSFET fabrication and reliability in mass production, as well as the rapidly increasing power dissipation may slow down the so far exponential scaling of complimentary MOSFETs (CMOS). Besides, with the ongoing search for new technological solutions vital for CMOS downscaling, developing conceptually new devices and architectures is becoming increasingly important. New nanoelectronic structures, such as carbon nanotubes, nanowires, and molecules, are considered to be the most prominent candidates for the post-CMOS era. Since conventional MOSFETs are already operating in the sub-100 nm range, new nanoelectronic devices are expected to complement and substitute some of the current CMOS functions after being integrated into CMOS technology.

Technology computer-aided design (TCAD) tools are used to assist in development and engineering at practically all stages ranging from process simulation to device and circuit optimization. The main purpose of TCAD is the technology-development related cost reduction which currently amounts to 35% and is expected to rise to 40%, according to ITRS [6]. Due to the aggressive downscaling of CMOS device feature sizes and newly emerging nanoelectronic devices, various shortcomings of presently applied TCAD tools appear. These tools are frequently based on semiclassical macroscopic transport models. From an engineering point of view, classical models like the drift–diffusion model, have enjoyed an amazing success due to their relative simplicity, numerical robustness, and the ability to perform two- and three-dimensional simulations on large unstructured grids [7]. Hot-carrier effects have motivated the development of higher-order transport models such as the hydrodynamic, energy-transport and six-moments models [8]. However, inaccuracies originate from the non-local nature of carrier propagation in ultra-scaled devices [9].

Non-local effects may be of classical or quantum-mechanical nature, depending on the underlying physics relevant to the transport process. Classical non-localities appear when the mean-free path is comparable to the device feature size. Quantum-mechanical non-local effects start to determine the transport properties when the device size is of the order of the De-Broglie electron wave length. Size quantization of carrier motion in inversion layers of MOSFETs and in ultra-scaled multi-gate devices as well as the tunneling current, including the gate leakage current, are the most important examples of quantum effects in MOSFETs.

Fig. 1 shows the hierarchy and mutual interrelation of models currently used for the description of current transport. Semiclassical transport models are based on the Boltzmann equation which includes scattering integrals describing realistic microscopic processes. These semiclassical models, augmented with quantum corrections, are still of great importance due to their relative computational simplicity, numerical stability, and an ability to provide reasonable quantitative results within seconds even for devices with gate length as short as 50 nm. A brief overview of the currently developed semiclassical transport models will be presented in Section 2.

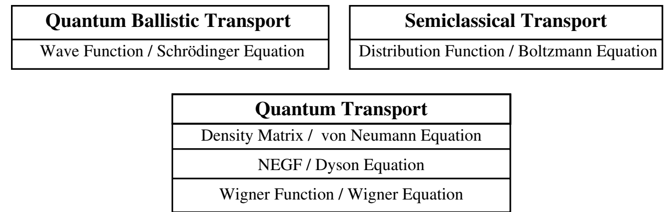

Fig. 1. Schematic classification of approaches used in semiconductor current transport modeling.

Quantum-ballistic transport models describe a coherent propagation of carriers. They are based on the solution of the Schrödinger equation for the wave function, supplemented with the corresponding boundary conditions. This approach is efficient and provides accurate results when carrier scattering is irrelevant and can be neglected. The method will be illustrated in Section 3 with an example of transport in carbon nanotubes [10].

Finally, dissipative quantum transport theory represents the most complete description of transport, which combines the coherent carrier motion between the scattering events with coherence (or phase) breaking due to carrier scattering. Different formalisms are currently used, based on the Dyson equation for the non-equilibrium Green's functions, the Liouville/Von-Neumann equation for the density matrix, or the Wigner transport equation. Section 4 deals with quantum transport characterized by both scattering and quantization. A conclusion will summarize the main findings and give directions for future research.

2. Semiclassical transport

After the ground-breaking work of Scharfetter and Gummel [11], who first proposed a robust discretization scheme for the drift–diffusion equation, computer programs like MINIMOS [12] and PISCES [13] played a pioneering role in numerical simulation of current transport properties of semiconductor devices. Since then, numerous transport models of increasing complexity have been introduced. The semiclassical transport description is based on the Boltzmann equation for the distribution of carriers $f(\mathbf{r}, \mathbf{k}, t)$ in the phase space. The Boltzmann equation includes carriers' scattering with phonons, impurities, interfaces, and other scattering sources through the corresponding collision integrals. Although the solution of the Boltzmann equation can be found numerically by means of Monte Carlo (MC) methods, TCAD models based on moments of the distribution function are highly desirable. Being computationally significantly less expensive than the MC method, these higher-order moments' methods provide a reasonable quantitative answer for devices as short as 50 nm within seconds. The fairly new six-moments

model [8] based on non-Maxwellian distribution takes naturally into account the hot-carrier effects such as avalanche generation, hot carrier induced gate currents, or hot-carrier diffusion, which typically take place in silicon-on-insulator (SOI) floating body MOSFETs. For the purpose of calibration the full-band MC method is often accepted, since it can precisely account for the various scattering processes [14].

Another important development of transport models is related to the MC methods for solving the Boltzmann equation. After the pioneering work of Kurosawa in 1966 [15], who was the first to apply the MC method to simulate carrier transport in semiconductors, the significantly improved MC method was successfully applied to transport description in a variety of semiconductors [16]. For electrons in silicon, the most thoroughly investigated case, it is believed that a satisfactory understanding of the band structure and of the basic scattering mechanisms has been achieved giving rise to a "standard model" [17]. Nowadays, an accurate MC evaluation of carrier transport properties in inversion layers is of primary importance for predicting performance of modern CMOS bulk devices. Due to the strong confinement of carriers in the inversion layer of bulk MOSFETs or due to the geometric confinement in multi-gate FETs the carrier motion is quantized in one or two confinement directions giving rise to the formation of subbands. One possibility to address the effect of quantum confinement on the electron concentration is to use an effective potential. This can be achieved by a convolution of the electrostatic potential with a Gaussian function, which leads to a smoothing of the original potential [18–20]. Another option is to use the self-consistent Poisson–Schrödinger-based quantum corrected potential [21,22], which suppresses the carrier concentration close to the interface, mimicking the real quantum-mechanical behavior. These approaches combine advantages of full-band structure and flexibility of scattering processes of three-dimensional classical MC simulations with the generality of material composition and transport peculiarities due to quantum confinement and may also address the strain effects.

The MC approach may incorporate the quantized carrier motion in the direction orthogonal to the current exactly. The quantum-mechanical motion of carriers in the confined direction is addressed by the self-consistent solution of the corresponding Schrödinger and Poisson equation, leading to the formation of subbands. The carrier motion within each subband may still be considered semi-classical and therefore can be well described by the corresponding Boltzmann equation written for the subband distribution function $f_n(\mathbf{r}, \mathbf{k}, t)$. Because of possible carrier transitions between different subbands due to scattering, the collision integrals on the right-hand-side of the Boltzmann equation should include the terms responsible for the intersubband scattering processes. The transport in the inversion layer of a MOSFET is finally described by a set of Boltzmann equations for every subband, coupled

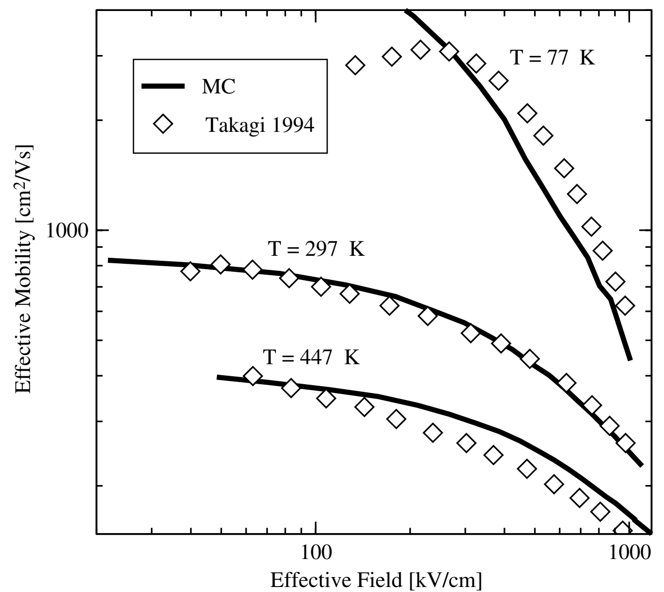

Fig. 2. Comparison of subband MC simulations with the experimental [23] universal mobility of surface layer in silicon. The deviation of the experimental mobility from simulations at low effective fields is due to Coulomb scattering not included in the MC simulations.

to each other via the intersubband scattering integrals. The set of the subband Boltzmann equations for $f_n(\mathbf{r}, \mathbf{k}, t)$ is conveniently solved by a MC method. This approach therefore combines the advantages of a quantum description in confinement direction with a semiclassical description in transport direction and represents a transition between semiclassical and quantum-mechanical pictures. An example of the simulation of the low-field surface mobility in inversion layers of silicon, when the transport in the current direction may be treated semiclassically, is shown in Fig. 2, together with the experimental "universal mobility" curve [23]. In order to reproduce the universal mobility curve, up to 40 unprimed and 20 primed subbands formed at a (100) silicon interface were taken into account, with realistic electron–phonon and surface roughness scattering included [24].

3. Quantum-ballistic transport

With the aggressive downscaling of MOSFET dimensions continuing, the classical description of carrier motion in transport direction is gradually losing its validity. When the characteristic scale of the potential variation along the channel is comparable to the De-Broglie wave length of a carrier, a TCAD transport model must include the quantum effects in transport direction. If scattering processes can be ignored and particle propagation in the device is coherent, the carrier motion is determined by the solution of the Schrödinger equation, supplemented with open boundary conditions. In order to determine the current density J, it is enough to know the transmission coefficient $TC(\mathscr{E})$ as well as the supply function $N(\mathscr{E}_x)$ from the electrodes [25]:

$$J = \frac{4\pi m^* q}{h^3} \int_{E_{\min}}^{E_{\max}} \mathrm{TC}(\mathscr{E}_x) N(\mathscr{E}_x)\, \mathrm{d}\mathscr{E}_x. \quad (1)$$

A similar approach can also be used to determine the gate leakage current [26]. The solution of the Schrödinger equation with open boundary conditions can be achieved by means of the quantum transmitting boundary method [27,28]. An established alternative framework for these calculations is the non-equilibrium Green's functions method [29] in its reduced coherent version. It is conveniently used for one-dimensional studies of resonant tunneling diodes or carbon nanotubes. Simulators accounting for a full two-dimensional solution of the open-boundary Schrödinger equation have been reported and applied to the simulation of 10 nm double-gate MOSFETs [30,31].

It may appear that in the quantum-ballistic case the determination of the full wave function as a solution of the Schrödinger equation is not necessary and the knowledge of the transmission coefficient is enough for the current calculations. In the contact block reduction method [32] the transmission function is fully determined by the reduced contact part of the full Green's function. However, the carrier concentration alters the electrostatic potential in the device via the Poisson equation. The carrier concentration is proportional to the square of the wave function, implying that the accurate determination of the transmission coefficient and therefore the current requires a self-consistent solution of the Schrödinger and Poisson equation simultaneously. For quasi-one-dimensional transport this can be achieved straightforwardly [4]. An example of the output characteristics simulated for an ultra-thin body double-gate MOSFET with a gate length L as short as 2.5 nm is shown in Fig. 3. Surprisingly, even such a small

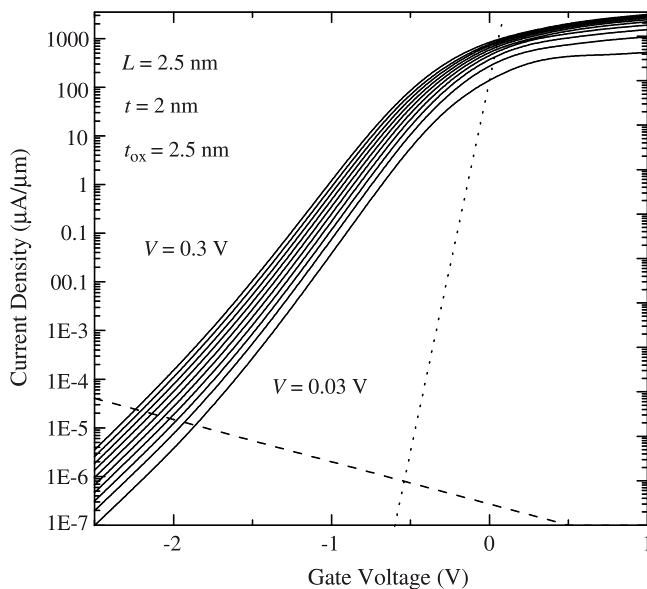

Fig. 3. Subthreshold characteristics for a double-gate MOSFET with silicon thickness of 2 nm, gate length of 2.5 nm and the oxide thickness of 2.5 nm [4]. The dotted line corresponds to the ideal 60 mV/decade subthreshold slope. The dashed line shows the leakage current.

device possesses an $I_{\mathrm{on}}/I_{\mathrm{off}}$ ratio sufficient for logic applications and displays a reasonable short-channel effect and acceptable DIBL, a conclusion recently reached from more detailed atomistic calculations [5]. It should be noted that the sensitivity to small MOSFET dimension variations, the control of doping as well as the whole manufacturing process development represent significant challenges for multi-gate MOSFETs with a gate length below 10 nm.

Self-consistent solution of the two- or three-dimensional Schrödinger equation together with the Poisson equation represents a significant computational challenge [30]. Two- and three-dimensional quantum-ballistic simulations can be performed by means of an approximate separation of the quantum motion in the confinement direction y from the motion along the current direction x by means of the following ansatz for the wave function $\Psi_n(y,x)$:

$$\Psi_n(y,x) = \sum_n \xi_n(x)\,\psi_n(y,x). \quad (2)$$

This method allows the independent solution of the Schrödinger equation for the subband wave function $\psi_n(y,x)$ at position x. Transport in the current direction is characterized by a system of one-dimensional Schrödinger equations with open boundary conditions for the wave functions $\xi_n(x)$. Each Schrödinger equation describes the transport inside the particular subband. Transport in each subband is independent from the one in other subbands, if the subband wave functions $\psi_n(y)$ do not depend on the position x in transport direction. The Schrödinger equations describing the transport in each subband are decoupled from each other, when the potential $U(x,y)$ in the device is the sum of two contributions, each depending either on y or x coordinate alone. In a general case when the subband wave functions depend on the position x in transport direction, the transport in the subbands n and m is coupled, with the coupling described by the Hamiltonian $\delta H_{nm}(x)$. However, when the intersubband coupling Hamiltonian $\delta H_{nm}(x)$ is small and may be neglected, transport in the subbands can still be considered as independent from each other. This approximation simplifies the calculations and reduces the computational effort significantly [33–36]. The coupling Hamiltonian is expected to be small if the dependence of the subband wave function on x is weak. An example where the subband decomposition turns out to be an excellent approximation is the quantum transport in ultra-scaled SOI MOSFETs [36]. In the opposite limit of abrupt junctions between contact reservoirs and the channel, the intersubband coupling is expected to be the strongest. However, even in this case the current value calculated self-consistently was found to be only 10% lower as compared to the calculations with neglected intersubband coupling [37]. More research is needed to clarify the situations.

The coherent quantum transport description is justified if the size of the channel region is shorter than the phase-breaking length. In carbon nanotubes, where elastic scattering can be ignored and inelastic scattering has little effect

on current [40], the current value can well be predicted within the quantum-ballistic approach [10]. Similar methods can be applied to describe the output characteristics of FinFETs in the ballistic approximation [41]. In silicon MOSFETs, however, the mean-free path in the area close to the potential maximum at 300 K is only a few nm [42], and the full quantum description which includes dissipative processes must be adopted to simulate MOSFETs with a gate length of around 10 nm. A consistent introduction of realistic scattering into simulators based on the coherent description alone creates outstanding computational difficulties ranging from a necessity to invert huge matrices in NEGF formalism [29] to calculations of non-local scattering rates in Pauli master equation approaches [43]. Besides the difficulties of introducing scattering into the simulators based on the coherent description, these simulators are often limited to specific geometries, grids and short-length scales, which makes their integration into modern engineering TCAD tools problematic. Nevertheless, these simulation approaches are necessary for the estimation of upper bounds of current transport at the quantum limit.

4. Dissipative quantum transport

The methods described so far are either based on the assumption of semiclassical or pure quantum mechanical ballistic transport. The former modeling approach has proven to be adequate to describe transport in previous generations of microelectronic devices. The latter one may be used for transport description when the carrier coherence length is larger than the device size. Recent studies show that even for devices with a channel length as short as 15 nm scattering will still play a significant role [44] and therefore determine the current, in accordance with estimations of the mean-free path in MOSFET structures [42]. The crossover from diffusive to ballistic transport in Si nanowire transistors occurs at approximately 2 nm [45], a much shorter distance than previously anticipated. An adequate transport model for ultra-scaled MOSFETS must therefore account for quantum-mechanical and dissipative effects simultaneously. In modern microelectronic devices quantum effects are usually dominant in a small active region connected to relatively large, heavily doped contact areas where the carrier dynamics is essentially classical. Therefore, modern TCAD simulators should also be able to incorporate both semiclassical and (dissipative) quantum-mechanical modeling approaches within the same formalism. To a certain extent, various quantum corrections can serve the purpose, as already discussed.

The non-equilibrium Green's functions method addresses the problem in the most consistent and complete way. Due to its completeness, the method is computationally complex and usually applied to one-dimensional problems [29] and for a restricted set of scattering mechanisms [46] only. The carbon nanotube (CNT) FET, which is widely considered to be a potential alternative to the conventional MOSFETs, represents a good example where the non-equilibrium

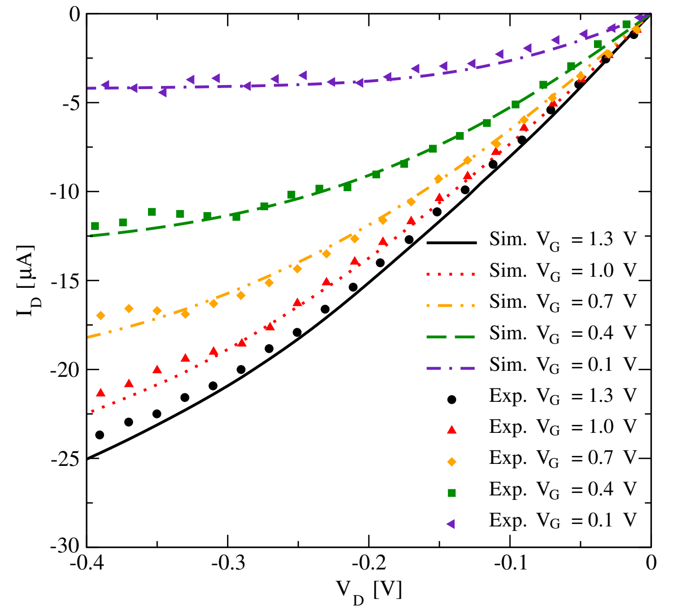

Fig. 4. Output characteristics of CNT-FET with ohmic contacts [38] compared to experimental data [39].

Green's functions method provides accurate results and is successfully used. In Fig. 4 simulated output characteristics of a CNT-FET with ohmic contacts [38] are compared to experimental data [39], showing good agreement.

An alternative approach which can handle both quantum-mechanical and dissipative scattering effects is based on the Wigner function formalism. Realistic scattering processes can be easily embedded into the Wigner equation via Boltzmann-like scattering integrals which turns out to be a good approximation. The Wigner function approach reduces to a semiclassical transport description in contacts providing an important advantage of a seamless treatment between classical and quantum-mechanical regions in device simulations [47].

The Wigner function is given by the density matrix in mixed representation [48,49] defined by the Wigner–Weyl transform

$$f_{\mathrm{w}}(\mathbf{r},\mathbf{k},t) = \int \rho\left(\mathbf{r}+\frac{\mathbf{s}}{2},\mathbf{r}-\frac{\mathbf{s}}{2},t\right)\exp(-\imath\mathbf{k}\cdot\mathbf{s})\mathrm{d}\mathbf{s}.$$

The kinetic equation for the Wigner function is similar to the Boltzmann equation:

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\cdot\nabla_r\right)f_{\mathrm{w}} = \int V_{\mathrm{w}}(\mathbf{r},\mathbf{k}'-\mathbf{k})f_{\mathrm{w}}(\mathbf{k}',\mathbf{r},t)\,\mathrm{d}\mathbf{k}' + \left(\frac{\partial f_{\mathrm{w}}}{\partial t}\right)_{\mathrm{coll}}. \quad (3)$$

The Wigner potential entering into the non-local operator in the right-hand side is defined as

$$V_{\mathrm{w}}(\mathbf{r},\mathbf{k}) = \frac{1}{\imath\hbar(2\pi)^3}\int\left(V\left(\mathbf{r}-\frac{\mathbf{s}}{2}\right)-V\left(\mathbf{r}+\frac{\mathbf{s}}{2}\right)\right)\exp(-\imath\mathbf{k}\cdot\mathbf{s})\,\mathrm{d}\mathbf{s}. \quad (4)$$

In case of slowly varying potentials the non-local potential term reduces to the classical force term. Following [50], one can introduce a spectral decomposition of the potential profile $V(x)$ into a slowly varying, classical component and a rapidly changing component treated quantum-mechanically.

$$V(\mathbf{r}) = V_{\mathrm{cl}}(\mathbf{r}) + V_{\mathrm{qm}}(\mathbf{r}). \tag{5}$$

The decomposition is conveniently carried out by applying a low-pass filter with a cut-off wave number $q_{\mathrm{c}} \ll \pi/\Delta x$, where Δx is a grid step size. This separation of the total potential into a classical and a quantum-mechanical contribution can improve the stability of a numerical solution method. The quantum-mechanical contribution may be moved into the right-hand side of the transport equation and can be interpreted as a quantum scattering integral:

$$\begin{aligned}&\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla_r - \frac{\mathrm{q}\nabla_r V_{\mathrm{cl}}(\mathbf{r})}{\hbar}\cdot\nabla_k\right) f_{\mathrm{w}} \\ &\quad = \int V_{\mathrm{qm}}(\mathbf{r},\mathbf{k}'-\mathbf{k}) f_{\mathrm{w}}(\mathbf{k}',\mathbf{r},t)\,\mathrm{d}\mathbf{k}' + \left(\frac{\partial f_{\mathrm{w}}}{\partial t}\right)_{\mathrm{coll}}.\end{aligned} \tag{6}$$

By applying the method of moments to (6), the quantum drift–diffusion or quantum hydrodynamic models can be derived [51]. These models are more convenient for the implementation in TCAD device simulators than a Schrödinger–Poisson solver which strongly depends on non-local quantities. However, it was reported that, while the carrier concentration in the inversion layer of a MOSFET is reproduced correctly, the method fails to account properly for tunneling currents [52].

The Wigner function formalism treats scattering and quantum-mechanical effects on equal footing through the corresponding scattering integrals. By analogy to the Monte Carlo methods used for the Boltzmann transport equation, it is tempting to try to solve the quantum Wigner transport Eq. (6) by means of the MC technique. Such a program was recently realized in [36,47,53]. However, since the kernel of the quantum scattering operator is not positively defined, the numerical weight of a particle trajectory increases rapidly, and the numerical stability of a trajectory-based Monte Carlo algorithm becomes a critical issue. A multiple trajectories method was suggested [47] to overcome this difficulty. In the algorithm developed, the problem of a growing statistical weight of a single trajectory is addressed by creating an increasing number of trajectories with constant weights, which may assume positive and negative values. Being formally equivalent to the former method, the algorithm allows the annihilation of particles with similar statistical properties, introducing a possibility to control the number of trajectories.

This method was recently applied to double-gate MOSFETs [36]. In the coherent mode, where scattering is turned off, a comparison to conventional Schrödinger solvers can be performed. In order to estimate the tunneling component of current the Wigner Monte Carlo simulations were carried out for a MOSFET with a gate length of 10 nm. Potential profiles were calculated self-consistently for a drain–source voltage of 0.4 V and a gate biases ranging from 0 V to 0.4 V using the device simulator MINIMOS-NT.

A good agreement between the two approaches is displayed in Fig. 5. The difference between quantum ballistic and semiclassical simulations is due to the additional contribution from electrons quantum-mechanically tunneling through the potential barrier. The increased concentration due to tunneling electrons is clearly seen in Fig. 6, where the normalized concentrations calculated semiclassically and from the Wigner equation are shown for two values of the gate voltage. The second moment of the Wigner function $f(r,\mathbf{p})$ and the semiclassical distribution function defined as

$$\langle E\rangle = \int \frac{p^2}{2m^*} f(r,\mathbf{p})\,\mathrm{d}\mathbf{p}, \tag{7}$$

is shown in Fig. 7. In the source electrode both quantum and classical simulations display the value $\langle E\rangle = 3/2kT$, which corresponds to the average energy of carriers injected with the Maxwellian distribution.

The difference in second moments of the classical ballistic and the Wigner functions is most pronounced in the area of potential maximum. Tunneling electrons formally possess a negative energy. This leads to a substantial negative contribution to (7) and an observed decrease of the energy moment of the Wigner function as compared to the classical case.

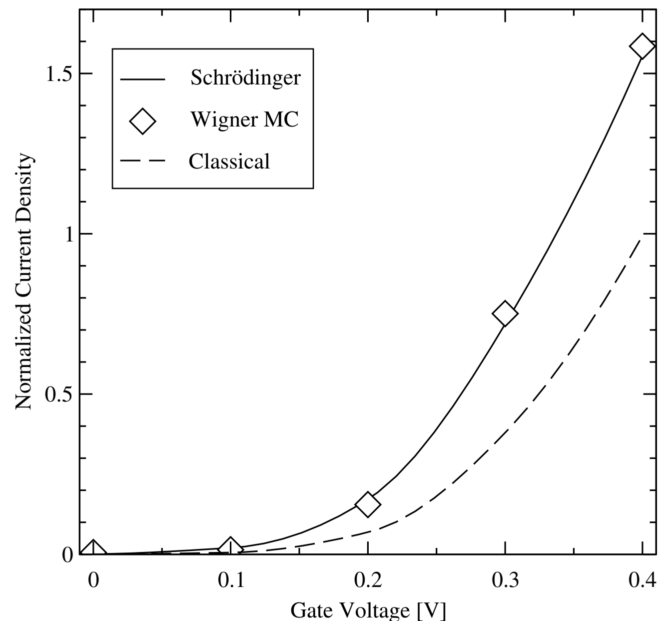

Fig. 5. Classical thermo-ionic current density normalized to its value at $V_{\mathrm{G}} = 0.4$ V and relative quantum-mechanical current density obtained with the Wigner Monte Carlo (open symbols). Symbol size shows the statistical uncertainty for the Wigner Monte Carlo simulations. Additional source-to-drain tunneling current component is clearly visible. Current densities found from the solution of the Schrödinger equation with open boundary conditions are also shown.

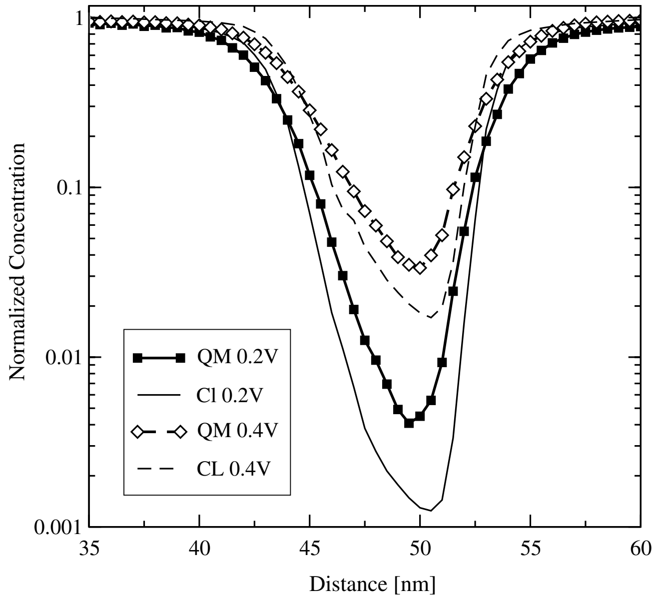

Fig. 6. The carrier concentrations in double-gate MOSFET, from classical (CL) and the Wigner (QM) Monte Carlo simulations, for two values of gate voltage.

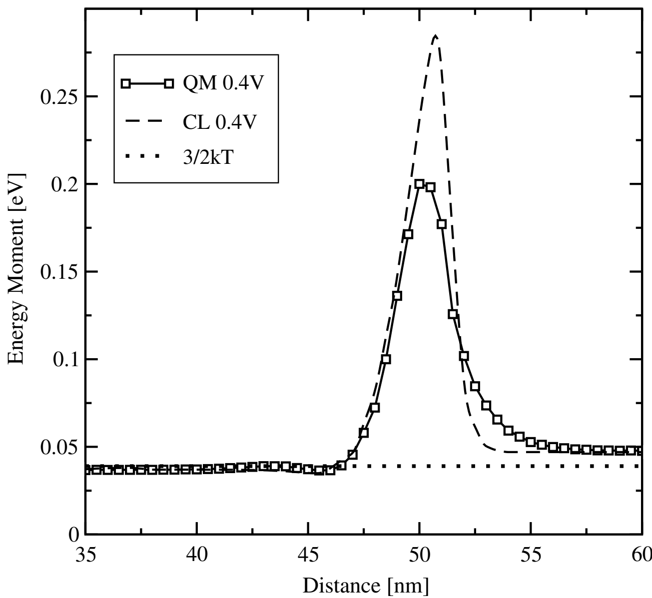

Fig. 7. The second moment, computed with the Wigner function (QM) and the classical ballistic distribution function (CL).

In the drain the average energies are again similar, although they show a higher value than those in the source. This is expected because of the ballistic nature of both classical and quantum transport considered. Indeed, electrons which reach the drain electrode by traveling from the channel to the drain electrode acquire sufficiently high energy due to the applied drain–source voltage. Due to absence of scattering, they preserve their high momenta leading to a significant contribution to the average energy throughout the drain area.

An important advantage of the Wigner Monte Carlo method is that it allows to include dissipative processes caused by scattering. It turns out that inclusion of scattering stabilizes the Wigner Monte Carlo simulations. The carrier concentration can be used to update the potential in the device by solving the Poisson equation. A superimposed iteration loop makes the Wigner–Poisson solver self-consistent. Examples of self-consistent potentials for n–i–n Si structures with an intrinsic region of length W ranging from 20 nm to 2.5 nm, as calculated with Wigner and classical Monte Carlo are shown in Fig. 8. The doping profile is assumed to increase gradually from the intrinsic channel to the highly doped contacts value over the same distance W. Electron–phonon and Coulomb scattering were included. As expected, for thick W the classical and quantum calculations yield similar results for the self-consistent potential. For $W = 2.5$ nm an extra space charge due to electrons tunneling under the barrier becomes important, which results in the potential barrier increase. Despite the potential barrier increase, the current in self-consistent Wigner simulations was approximately 20% higher compared to its classical value found by a self-consistent solution of the Boltzmann and the Poisson equations.

The Wigner function method gives accurate results not only for single-barrier devices, but can also be applied to purely quantum-mechanical systems such as resonant tunneling diodes [47]. A typical output characteristic of a GaAs resonant tunneling diode is shown in Fig. 9. Scattering with polar optical phonons as well as the Coulomb scattering in the contacts is considered. A region of negative differential resistance common to transport via a resonant level is clearly visible after the resonance peak at 250 mV applied

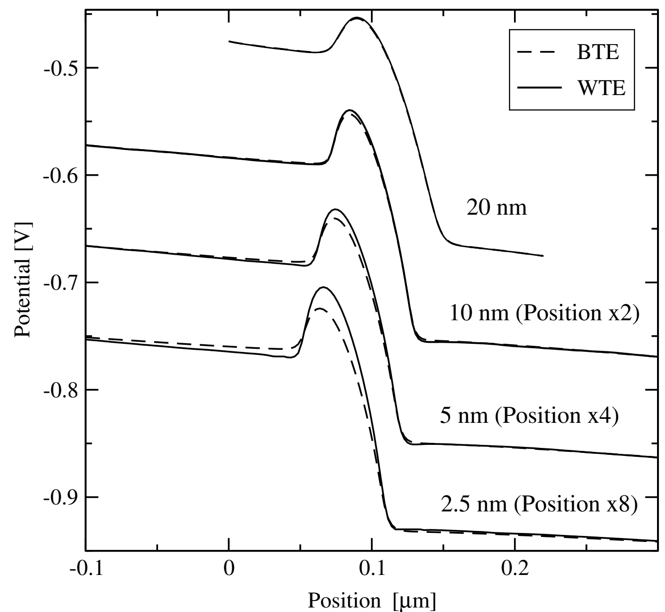

Fig. 8. Self-consistent potential profiles calculated for n–i–n structure with Wigner (solid lines) and Boltzmann (dashed lines) transport equations. For long n–i–n structures results are similar. For short n–i–n structures additional charge due to tunneling electrons results in significantly higher potential barrier.

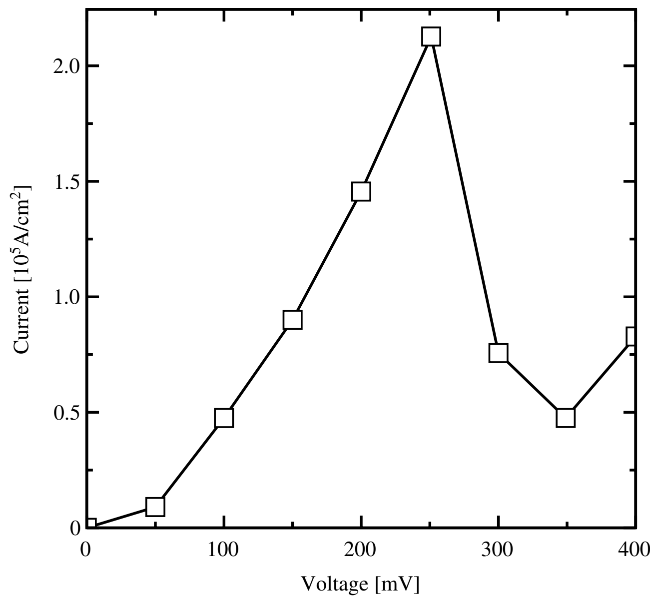

Fig. 9. Typical *IV* curve of resonant tunneling diode computed with Wigner MC self-consistently. The negative differential resistance after the peak is characteristic for resonant structures.

voltage. Self-consistent solution of the Wigner transport and Poisson equation is a mandatory for the correct determination of the resonance position due to charge accumulation at the cathode side of the resonant tunneling diode. A typical distribution of the concentration in resonance condition and off-resonance is shown in Fig. 10. The amount of charge localized in the potential well is much higher at resonance as compared to off-resonance conditions, in accordance with previous simulations [47]. This example demonstrates the importance of quantum-mechanical effects for simulations of properties of ultra-scaled devices. It also shows that space charge effects are of crucial importance for the accurate prediction of output characteristics of single- and double-barrier devices.

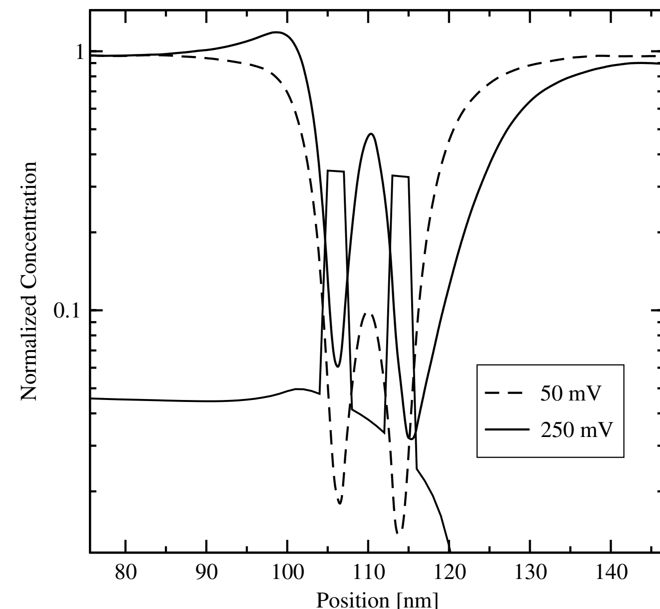

Fig. 10. Normalized electron concentration off-resonance (dashed line) and at resonance (solid line) in double barrier structure. Space charge accumulation seen at the anode side of resonant tunneling diode results in shift of *I*–*V* current resonant pick.

5. Conclusions

Well established classical TCAD tools are gradually losing their ability to predict accurately the characteristics of nanoscale devices, prompting for enhancement to meet the engineering demands. Classical models using higher moments are able to include the hot-carrier effects and can closely reproduce results of the full-band Monte Carlo. Relevant quantum corrections may be incorporated into the Monte Carlo simulators allowing to approximately account for some quantum phenomena.

Full quantum description is required for nanoscale devices. Contrary to the carbon nanotubes, where the transport properties can be predicted within the coherent picture, a dissipative quantum description may be required for transport calculations in ultra-scaled MOSFETs with gate lengths ranging below 10 nm. One option is the Wigner function approach which combines the advantages of quantum description with the accurate scattering models relevant for devices in the nanoscale range. All quantum-mechanical models must be adapted for engineering applications for which timely results are often more valuable than accurate analyses [54].

Acknowledgments

We gratefully acknowledge financial support from the Austrian Science Fund FWF, project P17285-N02, and the European Commission, project SINANO IST-506844.

References

[1] International technology roadmap for semiconductors, 2005 Edition. Available from: http://www.itrs.net/Common/2005ITRS/Home2005.htm.
[2] Doris B, Ieong M, Kanarsky T, Zhang Y, Roy RA, Documaci O, et al. Extreme scaling with ultra-thin Si channel MOSFETs. In: IEDM Tech Dig; 2002. p. 267–70.
[3] Iwai H. CMOS downsizing toward sub-10 nm. Solid-State Electron 2004;48(4):497–503.
[4] Sverdlov VA, Walls TJ, Likharev KK. Nanoscale silicon MOSFETs: a theoretical study. IEEE Trans Electron Dev 2003;50(9):1926–33.
[5] Risch L. Pushing CMOS beyond the Roadmap. In: Proc Eur Solid-State Dev Res Conf; 2005. p. 63–8.
[6] International technology roadmap for semiconductors, 2005 Edition, Table 124. Available from: http://www.itrs.net/Common/2005ITRS/Modeling2005.pdf.
[7] Selberherr S. Analysis and simulation of semiconductor devices. Springer; 1984.
[8] Grasser T, Tang T-W, Kosina H, Selberherr S. A review of hydrodynamic and energy-transport models for semiconductor device simulation. Proc IEEE 2003;91(2):251–74.
[9] Heinz FO, Bufler FM, Schenk A, Fichtner W. Quantum transport phenomena and their modeling. In: Symp Nano Dev Technol, Hsinchu, Taiwan; 2004. p. 2–8.

[10] Pourfath M, Gehring A, Ungersboeck E, Kosina H, Selberherr S, Cheong B-H, et al. Separated carrier injection control in carbon nanotube field-effect transistors. J Appl Phys 2005;97(10):1061031–33.
[11] Scharfetter DL, Gummel HK. Large-signal analysis of a silicon read diode oscillator. IEEE Trans Electron Dev 1969;16(1):64–77.
[12] Selberherr S, Fichtner W, Pötzl H. MINIMOS – a program package to facilitate MOS device design and analysis. In: Browne BT, Miller JJ, editors. Numerical analysis of semiconductor devices and integrated circuits, vol. I. Dublin: Boole Press; 1979. p. 275–9.
[13] Pinto MR. PISCES IIB. Stanford University; 1985.
[14] Fischetti MV, Laux SE. Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects. Phys Rev B 1988;38(14):9721–45.
[15] Kurosawa T. Monte Carlo calculation of hot electron problems. In: Proc Int Conf Phys Semicond; 1966. p. 424–6.
[16] Jacoboni C, Reggiani L. The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials. Rev Mod Phys 1983;55(3):645–705.
[17] Fischetti M, Laux A. Monte Carlo simulation of electron transport in Si: the first 20 years. In: 26th Eur Solid State Dev Res Conf; 1996. p. 813–20.
[18] Li Y, Tang T-W, Wang X. Modeling of quantum effects for ultrathin oxide MOS structures with an effective potential. IEEE Trans Nanotechnol 2002;1(4):238–42.
[19] Ahmed KZ, Kraus PA, Olsen C, Nouri F. On the evaluation of performance parameters of MOSFETs with alternative gate dielectrics. IEEE Trans Electron Dev 2003;50(12):2564–7.
[20] Palestri DEP, Eminente S, Fiegna C, Sangiorgi E, Selmi L. An improved semi-classical Monte-Carlo approach for nano-scale MOSFET simulation. Solid-State Electron 2005;49:727–32.
[21] Kathawala GA, Winstead B, Ravaioli U. Monte Carlo simulations of double-gate MOSFETs. IEEE Trans Electron Dev 2003;50(12): 2467–73.
[22] Fan X-F, Wang X, Winstead B, Register LF, Ravaioli U, Banerjee SK. MC simulation of strained-Si MOSFET with full-band structure and quantum corrections. IEEE Trans Electron Dev 2004;51(6): 962–70.
[23] Takagi S-I, Toriumi A, Iwase M, Tango H. On the universality of inversion layer mobility in Si MOSFET's: Part I – Effects of substrate impurity concentration. IEEE Trans Electron Dev 1994;41(12): 2357–62.
[24] Ungersböck E, Kosina H. The effect of degeneracy on electron transport in strained silicon inversion layer. In: Proc Int Conf Simul Semicond Processes Dev, Tokyo; 2005. p. 311–4.
[25] Tsu R, Esaki L. Tunneling in a finite superlattice. Appl Phys Lett 1973;22(11):562–4.
[26] Gehring A. Simulation of tunneling in semiconductor devices. Dissertation, Technische Universität Wien, 2003.
[27] Lent CS, Kirkner DJ. The quantum transmitting boundary method. J Appl Phys 1990;67(10):6353–9.
[28] Frensley WR. Numerical evaluation of resonant states. Superlattices Microstruct 1992;11(3):347–50.
[29] Lake R, Klimeck G, Bowen RC, Jovanovic D. Single and multiband modeling of quantum electron transport through layered semiconductor devices. J Appl Phys 1997;81(12):7845–69.
[30] Laux SE, Kumar A, Fischetti MV. Ballistic FET modeling using QDAME: quantum device analysis by modal evaluation. IEEE Trans Nanotechnol 2002;1(4):255–9.
[31] Sabathil M, Hackenbuchner S, Majewski JA, Zandler G, Vogl P. Towards fully quantum mechanical 3D device simulations. J Comput Electron 2002;1:81–5.
[32] Mamaluy D, Sabathil M, Vogl P. Efficient method for the calculation of ballistic quantum transport. J Appl Phys 2003;93(8): 4628–33.
[33] Heinz FO, Schenk A, Scholze A, Fichtner W. Full quantum simulation of silicon-on-insulator single-electron devices. J Comput Electron 2002;1(1):161–4.
[34] Curatola G, Fiori G, Iannaccone G. Modeling and simulation challenges for nanoscale MOSFETs in the ballistic limit. Solid-State Electron 2004;48(4):581–7.
[35] Wang J, Polizzi E, Ghosh A, Datta S, Lundstrom M. Theoretical investigation of surface roughness scattering in silicon nanowire transistor. J Appl Phys 2005;87:0431011–13.
[36] Sverdlov V, Gehring A, Kosina H, Selberherr S. Quantum transport in ultra-scaled double-gate MOSFETs: A Wigner function-based Monte Carlo approach. Solid-State Electron 2005;49(9):1510–5.
[37] Venugopal R, Goasguen S, Datta S, Lundstrom MS. Quantum mechanical analysis of channel access geometry and series resistance in nanoscale transistors. J Appl Phys 2003;95(1):292–305.
[38] Pourfath M, Kosina H, Cheong B, Park W, Selberherr S. Improving DC and AC characteristics of ohmic contact carbon nanotube field effect transistors. In: Proc Eur Solid-State Dev Res Conf, Grenouble; 2005. p. 541–4.
[39] Javey A, Guo J, Farmer D, Wang Q, Yenilmez E, Gordon R, et al. Self-aligned ballistic molecular transistors and electrically parallel nanotube arrays. Nano Lett 2004;4(7):1319–22.
[40] Guo J, Lundstrom M. Role of phonon scattering in carbon nanotube field-effect transistors. Appl Phys Lett 2005;86:193103-1–3-3.
[41] Shao X, Yu Z. Nanoscale FinFET simulations: a quasi-3D quantum mechanical model using NEGF. Solid-State Electron 2005;49: 1435–45.
[42] Jungemann C, Subba N, Goo J-S, Riccobene C, Xiang Q, Meinerzhagen B. Investigation of strained Si/SiGe devices by MC simulation. Solid-State Electron 2004;48(8):1417–22.
[43] Fischetti MV. Theory of electron transport in small semiconductor devices using the Pauli Master equation. J Appl Phys 1998;83(1): 270–91.
[44] Palestri SEP, Esseni D, Fiegna C, Sangiorgi E, Selmi L. Understanding quasi-ballistic transport in nano-MOSFETs: Part I – Scattering in the channel, and in the drain. IEEE Trans Electron Dev 2005;52(12): 2727–35.
[45] Gilbert MJ, Akis R, Ferry DK. Phonon-assisted ballistic to diffusive crossover in silicon nanowire transistors. J Appl Phys 2005;98(9): 094303-1–3-8.
[46] Svizhenko A, Anantram MP. Role of scattering in nanotransistors. IEEE Trans Electron Dev 2003;50:1459–66.
[47] Kosina H, Nedjalkov M, Selberherr S. A Monte Carlo method seamlessly linking quantum and classical transport calculations. J Comput Electron 2002;2(2–4):147–51.
[48] Wigner E. On the quantum correction for thermodynamic equilibrium. Phys Rev 1932;40:749–59.
[49] Kosina H, Nedjalkov M. Wigner function based device modeling. In: Rieth M, Schommers W, editors. Handbook of theoretical and computational nanotechnology. Springer; 2005.
[50] Gehring A, Kosina H. Wigner-function based simulation of quantum transport in scaled DG-MOSFETs using the Monte Carlo method. J Comput Electron 2005;4(1–2):67–70.
[51] Wettstein A, Schenk A, Fichtner W. Quantum device-simulation with the density-gradient model on unstructured grids. IEEE Trans Electron Dev 2001;48(2):279–84.
[52] Hoehr T, Schenk A, Wettstein A, Fichtner W. On density-gradient modeling of tunneling through insulators. In: Proc Int Conf Simul Semicond Processes Dev; 2002. p. 275–8.
[53] Shifren L, Ringhofer C, Ferry DK. A Wigner function-based quantum ensemble Monte Carlo study of a resonant tunneling diode. IEEE Trans Electron Dev 2003;50(3):769–73.
[54] Duane M. TCAD needs and applications from a users perspective. IEICE Trans Electron 1999;E82-C(6):976–82.