Wigner Monte Carlo Simulation: Particle Annihilation and Device Applications

Hans Kosina, Viktor Sverdlov, and Tibor Grasser
Institute for Microelectronics, TU Wien, Gußhausstraße 27–29/E360, A-1040 Wien, Austria
Phone: +43-1-58801-36013, Fax: +43-1-58801-36099, E-mail: hans.kosina@iue.tuwien.ac.at

***Abstract*— Coherent transport in mesoscopic devices is well described by the Schrödinger equation supplemented by open boundary conditions. When electronic devices are operated at room temperature, however, a realistic device model needs to include carrier scattering. In this work the kinetic equation for the Wigner function is employed as a model for quantum transport. Carrier scattering is treated in an approximate manner through a Boltzmann collision operator. A Monte Carlo technique for the solution of this kinetic equation has been developed, based on an interpretation of the Wigner potential operator as a generation term for numerical particles. Details on the algorithm for particle generation and subsequent particle annihilation are presented. Including a multi-valley semiconductor model and a self-consistent iteration scheme, the described Monte Carlo simulator can be used for routine device simulations. Applications to single barrier and double barrier structures are presented.**

I. Introduction

For FETs with gate lengths below 10 nm quantum effects such as direct source-to-drain tunneling become important and start affecting the device characteristics [1]. Recent studies show that scattering will still affect the current [2] and that the transition to ballistic transport appears at much shorter gate lengths than previously anticipated [3]. An accurate theory of MOSFETs near the scaling limit must therefore account for the interplay between coherent quantum effects and dissipative scattering effects. This mixed transport regime can suitably be treated by the Wigner equation. Early numerical solutions of the Wigner equation were obtained using finite difference methods, assuming simplified scattering models based on the relaxation time approximation [4]. However, for realistic device simulation more comprehensive scattering models are required. With the advent of Monte Carlo (MC) methods for the Wigner equation [5], [6] it became feasible to include the full Boltzmann collision operator. The development of MC methods for the Wigner equation, however, is hampered by the fact that, as opposed to the semi-classical case, the integral kernel is no longer positive. This so-called negative sign problem will lead to exponentially growing variances of the Markov Chain MC method. The Wigner potential operator can also be viewed as a generation term of positive and negative numerical particles. In this picture the sign problem shows up in the avalanche of numerical particles generated. A stable MC method can only be achieved by means of a suitable particle annihilation algorithm.

II. The Physical Model

Quantum transport is modeled by a time-independent, one-electron Wigner equation for a multi-valley semiconductor. The set of Wigner equations is coupled through the inter-valley phonon scattering terms.

$$\frac{1}{\hbar}\Big(\nabla_{\mathbf{k}}\epsilon_v(\mathbf{k})\cdot\nabla_{\mathbf{r}}+\mathbf{F}(\mathbf{r})\cdot\nabla_{\mathbf{k}}\Big)f_v(\mathbf{k},\mathbf{r}) = \sum_{v'}\int[1-f_v^0(\mathbf{k},\mathbf{r})]S_{vv'}(\mathbf{k},\mathbf{k}')f_{v'}(\mathbf{k}',\mathbf{r})\,d^3k' - \left(\sum_{v'}\int[1-f_{v'}^0(\mathbf{k},\mathbf{r})]S_{v'v}(\mathbf{k}',\mathbf{k})\,d^3k'\right)f_v(\mathbf{k},\mathbf{r}) + \int V_{\mathrm{w}}(\mathbf{k}-\mathbf{k}',\mathbf{r})f_v(\mathbf{k}',\mathbf{r})\,d^3k' \qquad \text{Silicon:}\quad v,\,v'=[100],\ [010],\ [001] \tag{1}$$

This equation determines the Wigner function f_v for valley v. A valley's energy dispersion relation $\epsilon_v(\mathbf{k})$ is assumed to be anisotropic and parabolic. Note that a non-parabolic $\epsilon(\mathbf{k})$ relation in the single-electron Hamiltonian would give a non-local diffusion term of the form $\int\hat{\epsilon}(\mathbf{k},\mathbf{r}-\mathbf{r}')f_v(\mathbf{k},\mathbf{r}')d^3r'$. Although it is straightforward to use a non-parabolic relation in (1), one should be aware that this would approximate the non-local diffusion term by a local one of the form $\nabla_{\mathbf{k}}\epsilon_v\cdot\nabla_{\mathbf{r}}f$.

A spectral decomposition of the potential profile $V(\mathbf{r})$ is applied [7]. The slowly varying component gives the classical force $\mathbf{F}$, whereas the rapidly varying component is taken into account through the Wigner potential V_{w}.

$$V(\mathbf{r}) = V_{\mathrm{cl}}(\mathbf{r}) + V_{\mathrm{qm}}(\mathbf{r}), \qquad \mathbf{F}(\mathbf{r}) = -\nabla V_{\mathrm{cl}}(\mathbf{r})$$

$$V_{\mathrm{w}}(\mathbf{q},\mathbf{r}) = \frac{1}{i\hbar(2\pi)^3}\int\Big[V_{\mathrm{qm}}\Big(\mathbf{r}+\frac{\mathbf{s}}{2}\Big)-V_{\mathrm{qm}}\Big(\mathbf{r}-\frac{\mathbf{s}}{2}\Big)\Big]\,\mathrm{e}^{-i\mathbf{q}\cdot\mathbf{r}}d^3s$$

In (1) scattering is treated semi-classically through a Boltzmann collision operator, where the transition rate $S_{vv'}(\mathbf{k},\mathbf{k}')$ from initial state $(v',\mathbf{k}')$ to final state $(v,\mathbf{k})$ is given by Fermi's golden rule. It should be noted that usage of the Boltzmann collision operator in the Wigner equation represents some ad hoc assumption. A rigorous treatment of electron-phonon scattering would require a frequency-dependent Wigner function, $f(\mathbf{k},\mathbf{r},\omega)$. It is related to the non-equilibrium Green's function $G^<$ by $G^<(\mathbf{r},\mathbf{k},\omega) = if(\mathbf{k},\mathbf{r},\omega)$ and can reasonably be

 1-4244-0404-5/06/$20.00

approximated as $f(\mathbf{k},\mathbf{r},\omega) = f_{\mathrm{w}}(\mathbf{k},\mathbf{r})A(\mathbf{r},\mathbf{k},\omega)$ [8]. To arrive at Fermi's golden rule the spectral function A is reduced to the Dirac δ-function.

Furthermore, in (1) the Pauli blocking factor the equilibrium Fermi function f_v^0 is used. The assumption of a Boltzmann collsion operator in (1) ensures that in the semiclassical regions, such as the highly doped contact regions, the conductivity is finite and that the mean energy increase due to degeneracy is taken into account.

III. Numerical Methods

A stationary MC method for solving (1) has been reported in [5]. The potential operator $\Theta[f_{\mathrm{w}}] = \int V_{\mathrm{w}}(\mathbf{k}-\mathbf{k}')f_{\mathrm{w}}(\mathbf{k}',\mathbf{r})d^3k'$ is interpreted as a generation term of numerical particles. The strict mass conservation property of this operator can be satisfied by the numerical particle model exactly if one generates the numerical particles only pair-wise, for instance, with statistical weights $+1$ and -1. As pointed out in [5], a suitable annihilation algorithm for numerical particles needs to be introduced in order to achieve a stable MC method. Since one can devise various algorithms for particle generation and, in particular, for particle annihilation, in the following the latest developments are described.

A. Particle Generation

A direct numerical representation of the Wigner potential $V_{\mathrm{w}}(\mathbf{q},\mathbf{r})$ would require the discretization of both momentum and space coordinates. The problem can be simplified by expressing the Wigner potential in terms of $\hat{V}(\mathbf{q})$, the Fourier transform of the potential $V_{\mathrm{qm}}(\mathbf{r})$. The potential operator can be rewritten as follows.

$$\Theta_{\mathrm{w}}[f_{\mathrm{w}}](\mathbf{k},\mathbf{r}) = \frac{1}{(2\pi)^3\hbar}\int |\hat{V}(\mathbf{q})|\,\sin\left[\varphi(\mathbf{q})+\mathbf{q}\cdot\mathbf{r}\right] \times\left(f_{\mathrm{w}}\left(\mathbf{k}-\frac{\mathbf{q}}{2},\mathbf{r},t\right)-f_{\mathrm{w}}\left(\mathbf{k}+\frac{\mathbf{q}}{2},\mathbf{r},t\right)\right)d^3q \quad (2)$$

An advantage of this formulation is that no discretization of the spatial variable $\mathbf{r}$ is needed. The expression can be evaluated at the actual position $\mathbf{r}$ of a particle. Only the momentum variable $\mathbf{q}$ needs to be discretized in order to numerically represent $|\hat{V}|$, the modulus, and φ, the phase of $\hat{V}$.

The structure of (2) suggests the usage of a rejection technique. As a normalization quantity one obtains an upper limit for the pair generation rate.

$$\gamma_{\max} = \frac{1}{(2\pi)^3\hbar}\int |\hat{V}(\mathbf{q})|\,d^3q \quad (3)$$

At a rate of $\gamma_{\max}$ the free flight of a particle is interrupted to check for particle pair-generation. From the distribution $|\hat{V}(\mathbf{q})|$ one generates randomly the momentum transfer $\mathbf{q}$. Then the sine function is evaluated at the actual particle position $\mathbf{r}$ as $s = \sin\left[\varphi(\mathbf{q})+\mathbf{q}\cdot\mathbf{r}\right]$. With probability $|s|$ the pair-generation event is accepted, otherwise a self-scattering event is performed. In the former case, two particle states are generated with momenta $\mathbf{k}_1 = \mathbf{k}-\mathbf{q}/2$ and $\mathbf{k}_2 = \mathbf{k}+\mathbf{q}/2$ and statistical weights $w_1 = w_0\mathrm{sign(s)}$ and $w_2 = -w_1$, respectively, where w_0 is the statistical weight of the initial particle. Since (2) is local in real space, the particle pair is generated at the position $\mathbf{r}$ of the initial particle.

B. Particle Annihilation

Different variants of the single-particle MC method outlined in [5] can be devised. The variant discussed below is constructed such that current is conserved exactly. The only input parameter required is the ratio of negative and positive trajectories, which makes the algorithm easy to control. The idea is that from the trajectory tree generated by a particle injected at the contact only one branch is actually traced.

For steady state problems considered here a phase space mesh can be utilized, on which numerical particles are temporarily stored. After each generation event one has to deal with three particle states, namely the initial state $\mathbf{k}$ and the two generated states, $\mathbf{k}_1$ and $\mathbf{k}_2$. In a first step all three particles are stored on the annihilation mesh, that is, the statistical weight of each particle is added to a counter associated with the mesh element. Then one has to decide which of the three states is used to continue the trajectory. One may choose the weight of the particle continuing the trajectory to be of the same sign as the incoming one (Fig. 1). In this way the statistical weight along one trajectory does not change, which results in exact current conservation. If the initial state has a positive statistical weight, out of the three mesh elements one selects that with the largest stored weight. Continuing from that element will reduce the weight of the element. Conversely, a negative trajectory is to be continued from the element with the smallest stored weight. A certain fraction of negative trajectories needs to be constructed in order to resolve the negative parts of the Wigner function. This rule for selecting the continuing particle is an attempt to minimize the weights stored in the three elements after each pair-generation event. The repeated execution of this rule in the MC main loop results in a minimization of the stored weight on the whole annihilation mesh. Particle annihilation takes place when positive and negative particles are alternately stored in the same mesh element. Note that because of the mass conservation property of the transport equation and of the associated particle model, no net-charge can build up on the annihilation mesh. The weights stored on the mesh sum up

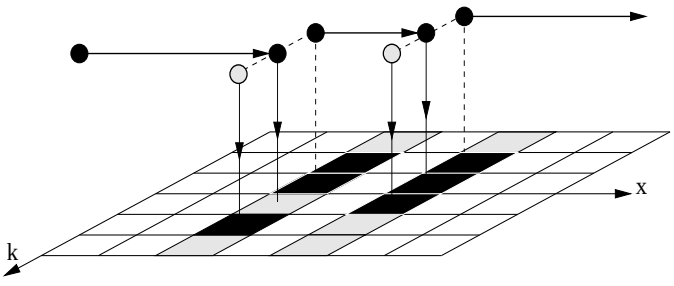

Fig. 1. The particle annihilation strategy attempts to minimize the weights stored in the mesh elements. The weights of the initial and continuing particle have the same sign to ensure current continuity. Particles and mesh elements carrying a positive weight are in black, the ones carrying a negative weight are in grey.

to zero. The local weights on the mesh have to be kept small, as they are a measure for the numerical error of the method. This can be controlled by the fraction of negative trajectories, which has to be specified by the user.

Considering Fig. 1, one can develop also the following notion of the algorithm. Positive and negative particles are sent through the device and interact with the annihilation mesh. A positive particle is likely to recombine in a phase-space region where the weight stored on the mesh is negative. This means that a positive particle is likely to propagate in those regions where the stored weight is positive. It is unlikely to recombine there, because this would result in an increase in local weight, which would contradict the local minimization principle. For the same reason, a negative particle will be attracted by regions with negative stored weight. Although the transition probabilities used to propagate the particles are the same for positive and negative particles, the interaction with the annihilation mesh causes the trajectories for positive and negative particles to be systematically different.

An annihilation mesh is introduced for each valley-type. For the three pairs of X-valleys of Si three meshes are required. The meshes are defined in the three-dimensional phase-space, spanned by one spatial and two momentum coordinates.

C. Coupling to Poisson Equation

A self-consistent iteration scheme between Wigner MC and the Poisson equation is implemented. The adopted scheme, which is similar to the Gummel iteration scheme for the basic semiconductor equations [9], is commonly used in classical one-particle MC simulations [10]. Fig. 2 shows the iteration history of the current through a Si *n-i-n* diode for different widths of the intrinsic region. Currents computed using Wigner and classical MC show similar convergence behavior.

IV. Results and Discussion

The described MC method can be used for routine device simulations. For the purpose of verification, the first example assumes a frozen potential profile from a 10 nm gate length double-gate MOSFET. Fig. 3 compares the quantum ballistic currents as obtained from a collision-less Wigner MC simulation and from a numerical Schrödinger solver. Good agreement is observed. The quantum ballistic current is higher than the classical ballistic current due to an additional contribution from carriers tunneling through the potential barrier.

To study the effects of scattering and tunneling on the device characteristics we consider Si *n-i-n* diodes with the length W of the intrinsic region ranging from 20 nm down to 2.5 nm. The doping profile is assumed to increase gradually from the intrinsic region to the highly doped contact region over the same distance W. Three transport models are compared: Wigner equation and Boltzmann equation with electron-phonon and ionized-impurity scattering included, yielding currents I_{WIG} and I_{BTE}, respectively. The Wigner equation without scattering inside the intrinsic and transition regions gives the current I_{COH} (coherent). Fig. 4 shows that the effect of scattering

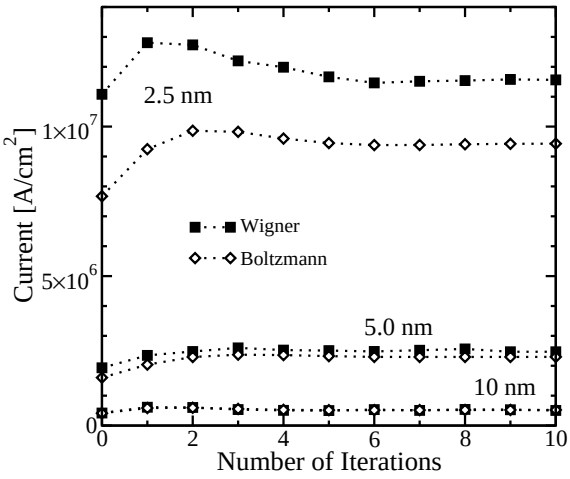

Fig. 2. Current through Si *n-i-n* diodes as a function of the number of self-consistent iterations with the width of the intrinsic region as a parameter.

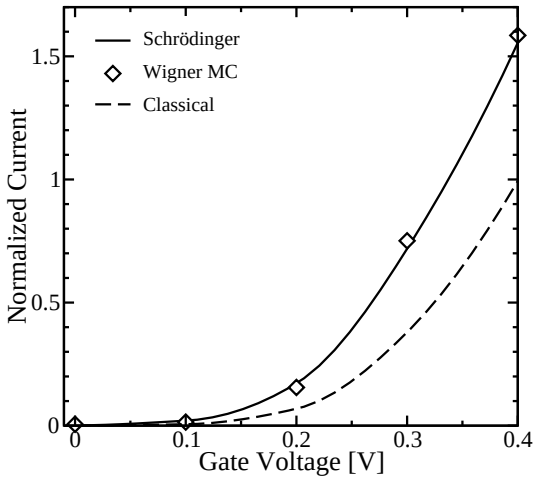

Fig. 3. Normalized ballistic currents calculated classically and quantum mechanically. Results from Wigner MC and the Schrödinger solver are in good agreement. The potential profile is obtained from a device simulation of a 10 nm gate length DG MOSFET.

reflected in the difference $I_{\mathrm{COH}} - I_{\mathrm{WIG}}$ decreases with decreasing device length. However, even for W = 2.5 nm the relative difference in the currents is still of the order of 25%, indicating that scattering cannot be neglected. Also shown is the current difference due to tunneling, $I_{\mathrm{WIG}} - I_{\mathrm{BTE}}$. Clearly, this current component rises with reduced barrier width.

The next example shows results of self-consistent Wigner-Poisson simulations of a double-barrier tunneling structure. A GaAs/AlGaAs resonant tunneling diode (RTD) is investigated, assuming a barrier height of $E_b = 0.3$ eV, a barrier width of 3 nm, and a well width of 5 nm [5], [11]. Polar optical phonon, acoustic deformation potential and ionized impurity scattering are included. Fig. 5 shows the effect of degeneracy, which is introduced in the simulation by the approximated Pauli blocking factors in (1) and through the boundary distribution at the contacts, on the current-voltage characteristics.

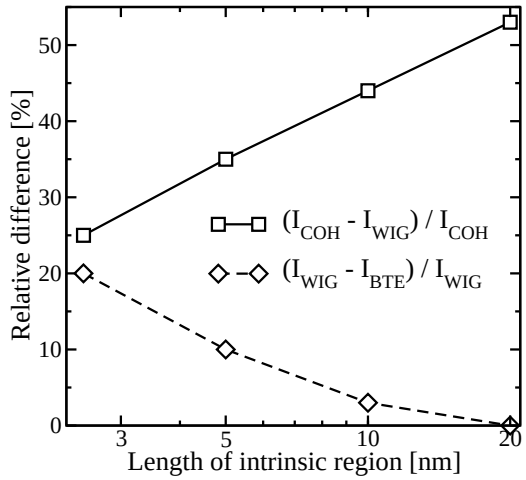

Fig. 4. Relative difference between currents of an *n-i-n* diode calculated using different transport models: Wigner MC with and without scattering in the intrinsic region (squares); Wigner MC and classical MC (diamonds).

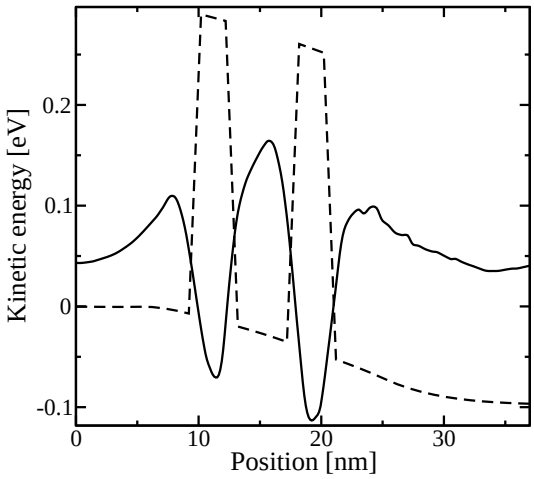

Fig. 6. Mean kinetic energy of electrons in a resonant tunneling diode calculated from (5). In the tunneling barriers the mean kinetic energy is negative.

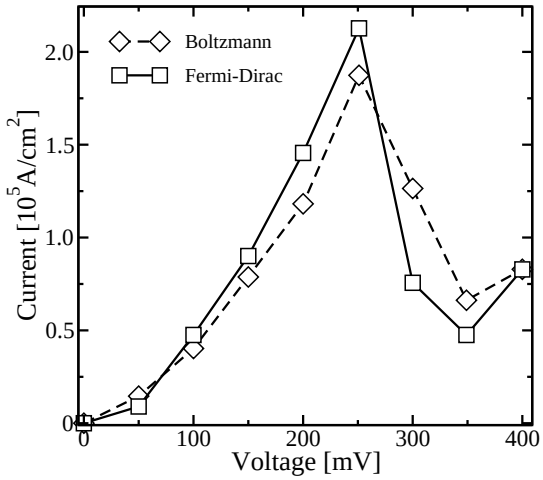

Fig. 5. I-V characteristics of a GaAs resonant tunneling diode at 300 K with scattering from PO phonons, acoustic phonons and ionized impurities included.

Fig. 6 shows the kinetic energy density of electrons in the RTD. In terms of wave functions ψ_i, the eigen energies E_i, and the probabilities p_i, the kinetic energy density is defined as follows.

$$w(\mathbf{r}) = \sum_i p_i (E_i - V(\mathbf{r})) |\Psi_i(\mathbf{r})|^2 \qquad (4)$$

This density can become negative in tunneling regions where the energy of one or more states is below the band edge, $E_i < V(\mathbf{r})$. Transformation of (4) into the Wigner representation gives [12]

$$w(\mathbf{r}) = \frac{1}{(2\pi)^3} \int \frac{\hbar^2}{2m^*} \left(|\mathbf{k}|^2 - \frac{1}{4} \nabla_{\mathbf{r}}^2 \right) f_{\mathrm{w}}(\mathbf{k}, \mathbf{r}, t)\, d^3k\,. \qquad (5)$$

Note that the kinetic energy density is not simply given by the second moment of the Wigner function. A correction proportional to the second derivative of the electron concentration is required.

V. Conclusion

A Monte Carlo simulator performing a self-consistent numerical solution of the Wigner equation has been presented. Details of the algorithms for generation and annihilation of numerical particles have been described. The quantum MC method turns gradually into the classical MC method when the potential profile becomes smoother. Therefore, the simulation method can be used, for instance, to study the gradual emergence of quantum effects when a device structure is scaled down.

Acknowledgment

This work has been partly supported by the Austrian Science Fund, project I79-N16.

References

[1] J. Wang and M. Lundstrom, in *Int.Electron Devices Meeting* (IEEE, 2002), pp. 29.2.1–29.2.4.
[2] P. Palestri, D. Esseni, S. Eminente, C. Fiegna, E. Sangiorgi, and L. Selmi, IEEE Trans.Electron Devices **52**, 2727 (2005).
[3] M. Gilbert, R. Akis, and D. Ferry, J.Appl.Phys. **98**, 094303 (2005).
[4] W. Frensley, Rev.Mod.Phys. **62**, 745 (1990).
[5] H. Kosina, M. Nedjalkov, and S. Selberherr, Journal of Compuational Electronics **2**, 147 (2003).
[6] L. Shifren, C. Ringhofer, and D. Ferry, Phys.Lett.A **306**, 332 (2003).
[7] A. Gehring and H. Kosina, Journal of Compuational Electronics **4**, 67 (2005).
[8] W. Hänsch, *The Drift-Diffusion Equation and Its Applications in MOSFET Modeling*, *Computational Microelectronics* (Springer Verlag, 1991).
[9] S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Springer, 1984).
[10] F. Venturi, R. Smith, E. Sangiorgi, M. Pinto, and B. Ricco, IEEE Trans.Computer-Aided Design **8**, 360 (1989).
[11] L. Shifren and D. Ferry, IEEE Trans.Electron Devices **50**, 769 (2003).
[12] H. Kosina and M. Nedjalkov, in *Handbook of Theoretical and Computational Nanotechnology*, edited by M. Rieth and W. Schommers (American Scientific Publishers, 2006), Vol. 10, Chap. Wigner Function Based Device Modeling, pp. 731–763, (in print).