Effect of boundary treatments on quantum transport current in the Green’s function and Wigner distribution methods for a nano-scale DG-MOSFET

Haiyan Jiang a, b, Wei Cai b, *

a Department of Applied Mathematics, Beijing Institute of Technology, Beijing 100081, China
b Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223-0001, United States

ABSTRACT

In this paper, we conduct a study of quantum transport models for a two-dimensional nano-size double gate (DG) MOSFET using two approaches: non-equilibrium Green’s function (NEGF) and Wigner distribution. Both methods are implemented in the framework of the mode space methodology where the electron confinements below the gates are pre-calculated to produce subbands along the vertical direction of the device while the transport along the horizontal channel direction is described by either approach. Each approach handles the open quantum system along the transport direction in a different manner. The NEGF treats the open boundaries with boundary self-energy defined by a Dirichlet to Neumann mapping, which ensures non-reflection at the device boundaries for electron waves leaving the quantum device active region. On the other hand, the Wigner equation method imposes an inflow boundary treatment for the Wigner distribution, which in contrast ensures non-reflection at the boundaries for free electron waves entering the device active region. In both cases the space-charge effect is accounted for by a self-consistent coupling with a Poisson equation. Our goals are to study how the device boundaries are treated in both transport models affects the current calculations, and to investigate the performance of both approaches in modeling the DG-MOSFET. Numerical results show mostly consistent quantum transport characteristics of the DG-MOSFET using both methods, though with higher transport current for the Wigner equation method, and also provide the current–voltage (I–V) curve dependence on various physical parameters such as the gate voltage and the oxide thickness.

1. Introduction

With fast development of semiconductor technologies, MOSFET dimensions are scaled down continuously. Gate and channel lengths are considered as the characteristic size of a MOSFET. The classical Boltzmann equation can accurately describe the drift-diffusion transport of charge carriers when the characteristic size is much larger than the mean free path of the carriers. However, quantum transport models should be used to address quantum effects once the characteristic size becomes much smaller than the mean free path [1]. The quantum transport models from the Schrödinger wave function can be implemented with either the formulation of the Non-equilibrium Green’s function (NEGF) [2] or that of the Wigner distribution function in a phase space [3]. Many simulations have been done on quantum devices such as the RTD (Resonant Tunneling Diode), bulk MOSFET, SOI MOSFET, and double gate (DG)-MOSFET [1, 4–7]. Especially, the DG-MOSFET with symmetric oxide layers and gates is a promising new device for better and more effective control of short channel effects.
A comprehensive description of quantum transport in nano-size MOSFETs is offered by the solution of NEGF, coupled with a Poisson equation self-consistently. However, the computational cost for solving the full NEGF is prohibitive; thus simplified models are usually employed to reduce the computational cost [6]. For thin body DG-MOSFETs, for which the confinement effect of the gates is strong, we could approximately decouple the solution of two-dimensional (2D) Schrödinger wave functions into two 1D problems [6,7]. Such an approximation is the basis of the mode space method where the electron confinement below the gates are pre-calculated to produce subbands along the vertical direction of the device while the channel transport is described by 1D Schrödinger equations with a subband energy profile along the transport direction. In many cases, one more approximation is done by assuming the eigenfunctions in the confinement direction do not change along the transport direction. Thus, those 1D Schrödinger equations for all subbands become decoupled, significantly reducing the total cost. Venugopal et al. analyzed the effectiveness of the mode space method by comparing it with the full real physical space discretization of the 2D Schrödinger equation for an ultra small DG-MOSFET [6]. It is concluded that the mode space method costs much less due to the fact that only a few subbands with lower energies need to be considered. The mode space method has also been used to compare the ballistic transport and the scattering transport of ultra thin body DG-MOSFETs [7].

In addition to the NEGF, a kinetic model can be derived using the Wigner distribution function [5,8] in the position-momentum phase space. The Wigner equation was first introduced by Wigner in 1932 by adding a correction term to the Boltzmann equation for a low temperature case [3]. Numerical methods for both NEGF and Wigner equations have attracted much attention recently due to the need of simulating quantum transport with computers [9,10]. The NEGF describes transport in a quantum open system using boundary self-energies to account for the effect of contacts to the device [20]. Within the NEGF formalism, a detailed treatment of the various scattering processes is possible [11]. On the other hand, the Wigner formalism has found many advantages for theoretical analysis of quantum transport. The Wigner function is an electron quasi-distribution in the phase space, which can model ideal contacts by separating incoming and outgoing components of the distribution at the boundaries. This phase space description is similar to classical distributions and allows us to incorporate a Boltzmann type collision term to explicitly deal with the electron scattering from ionized impurities, acoustic phonons, and surface roughness at the Si/SiO₂ interface [5,12]. The Wigner equation can be solved by a probability method (Monte Carlo) [13,14] and deterministic methods (Finite Difference method, Spectral method, etc) [1,4,5,15,16]. Both steady-state and transient Wigner equations have been solved by the finite difference method to analyze the transport character of RTD [4]. The mode space method with 1D Wigner equations along the channel direction has been applied to determine the current through a nano-size SOI MOSFET and analyze the effect of the channel size to the current character. Scattering effects due to impurity, acoustic phonon, and surface roughness at the boundary between the silicon layer and the oxide layer are also considered in the Wigner-mode space combination [1,5].

In this paper, we will study both NEGF and Wigner function methods for quantum transport along the channel direction and investigate the different manners the device boundary conditions are treated and their effects on the transport current calculations for a nano-scale DG-MOSFET. As the Wigner equation is a reformulation of the Schrödinger equation by a Weyl transform and the Fourier transform, the NEGF and the Wigner equation descriptions are in principle equivalent. However, different treatments of boundary conditions and scatterings produce different levels of approximation accuracy and computational efficiency. In the case of NEGF, the contact boundaries are treated by self-energy terms which are basically Dirichlet to Neumann mappings for the Green's function on the boundaries [7,17,20]. The boundary conditions are so designed to observe the causality of the system through an outgoing radiation condition; as a result, electron waves leaving the active device regions will not be reflected at the boundaries. On the other hand, inflow boundary conditions are posed for the Wigner distribution such that free electrons entering the device region will not be reflected at the boundaries, and the Wigner distribution effectively assumes Fermi–Dirac equilibrium distributions of the electrons in the contacts at the appropriate Fermi levels. The difference of imposing non-reflecting properties in the NEGF and the Wigner distribution methods will have effects on the transport current calculated by either method, as shown in our simulations of a nano-scale DG-MOSFET.

The paper is organized as follows. In Section 2, we first introduce the mode space method for 2D Schrödinger equations and the concept of subband, then the relevant formulas for density functions and Landauer formulation for the current using transmission probability coefficients. Section 3 describes the method of the NEGF and its treatment of quantum boundary using the self-energy to ensure the non-reflection at the device boundaries for electron waves leaving the quantum device active region, and most importantly, the relation between the transmission probability coefficient and the NEGF, and the subsequent current formula using the NEGF. In Section 4, the Wigner function method is introduced, and the inflow boundary condition for the Wigner distribution, which ensures the non-reflection at the boundaries for free electron waves entering the device active region, is elaborated. Section 5 contains the numerical simulations with both the NEGF and the Wigner equations in the channel direction in the framework of the mode space method for a nano-scale DG-MOSFET. Finally, a conclusion and a discussion are given in Section 6.

2. Current formula for transport in a DG-MOSFET in mode space methods

A DG-MOSFET has a structure shown in Fig. 1 with a silicon layer sandwiched by two symmetric oxide layers. The source and the drain are doped heavily, while the body (L) is made intrinsic to approximate the ballistic limit [7]. As the scale of the
device is very large in the y direction, the static potential \( V \) is assumed invariant along the y direction and determined by a 2D Poisson equation
\[
-\nabla \cdot (\varepsilon(x, z) \nabla V(x, z)) = e(-\rho(x, z) + N_d(x, z)),
\]
where \( e \) is the electron charge, \( \varepsilon(x, z) \) is the dielectric, \( \rho(x, z) \) is the electron density, and \( N_d(x, z) \) is the doping density function, respectively. The boundary condition is
\[
\begin{aligned}
V(x, z) &= V_g - (W^m_{ij} - W^o_{ij}), & (x, z) \in EF, \ GH, \\
\frac{\partial V(x, z)}{\partial \eta} &= 0, & (x, z) \in AB, \ BG, \ HC, \ CD, \ DF, \ EA,
\end{aligned}
\]
where \( \eta \) is the normal direction of the boundary, and \( W^m_{ij} = 4.188 \) is the work function of the metal gate, and \( W^o_{ij} = 4.05 \) is that of the silicon. The Poisson equation is solved in the square domain ‘ABCD’ which includes the silicon layer and the oxide layers. On EF and GH, Ohmic contacts are imposed with a gate voltage \( V_g \). Meanwhile, insulator contacts are assumed on boundaries of the oxide layers and the floating boundary condition is imposed on the contacts to maintain macroscopic space-charge neutrality at the source (drain) contact irrespective of the biasing condition. The electron density function \( \rho(x, z) \) has different representations in classical or quantum descriptions of the charge carrier transport. For a large scale device, charge carriers can be considered as classical particles for which the Newton’s motion law can be used to describe the dynamics of the distribution of charge carriers in the Boltzmann equation.

As the device shrinks, a quantum transport model needs to be considered [10]. For 2D devices, we consider the 2D Schrödinger equation
\[
H \Psi(x, z) = E \Psi(x, z),
\]
where \( H \) is the Hamiltonian operator given by
\[
H = -\frac{\hbar^2}{2m_x} \frac{\partial^2}{\partial x^2} - \frac{\hbar^2}{2m_y} \frac{\partial^2}{\partial y^2} - \frac{\hbar^2}{2m_z} \frac{\partial^2}{\partial z^2} + eV(x, z).
\]
Here \( \hbar \) is the reduced Plank constant, and \( m_x, m_y, m_z \) are the effective mass in the x, y, z directions, respectively. In fact, there are three valleys in the band-structure of the silicon and each valley has a different effective mass. The effective mass can be denoted by a vector \( m_x = (0.19, 0.98, 0.19)m_0, m_y = (0.19, 0.19, 0.98)m_0, m_z = (0.98, 0.19, 0.19)m_0 \), where \( m_0 \) is the mass of free electron in vacuum. The total density and the total current are calculated by a summation over the three valleys which loops all conduction band minima of a given material [21]. For simplicity, the following formula for the density or current is given for one valley. For a non-equilibrium state over an infinite domain, a finite domain problem for the active device region can be solved either by the Green’s function method or by the QTBM (Quantum-Transport-Boundary-Method) [18,19]. Either method has its limitations: the Green’s function method incurs a large computational cost while the QTBM can not handle the whole contact effect.

2.1. A mode space method

As the solution of the full 2D Schrödinger equation is expensive in computing cost, reduction to simpler models is usually preferred, and a popular approach is the mode space method. Though being approximation in nature, the mode space method has been shown to give solutions reasonably close to those by the original full 2D solutions for ultra thin body DG-MOSFETs [6]. Within the mode space framework, the following 2D wave function \( \Psi(x, z) \) is expanded with eigenfunctions \( \Theta_m(x, z) \) for the confinement direction,
\[
\Psi(x, z) = \sum_m \Theta_m(x, z) \phi_m(x),
\]
where \( \Theta_m(x, z) \) is the eigenfunction in the \( z \) direction, satisfying the following 1D Schrödinger equation

\[
-\frac{\hbar^2}{2m_z} \frac{\partial^2 \Theta_m(x, z)}{\partial z^2} + eV(x, z)\Theta_m(x, z) = E_m(x)\Theta_m(x, z) \quad \text{for a fixed} \ x.
\]  

(5)

If the electron penetration into the oxide is ignored, zero Dirichlet boundary conditions can be used,

\[
\Theta_m(x, 0) = 0, \quad \Theta_m(x, T_{si}) = 0,
\]

where \( T_{si} \) is the thickness of the silicon layer. Plugging (4) into the 2D Schrödinger Eq. (2) and assuming further that the eigenfunctions in the confinement direction \( z \) do not depend on \( x \), i.e.,

\[
\frac{\partial \Theta_m(x, z)}{\partial x} = 0,
\]

at the \( m \)th subband \( E_m(x) \), we arrive at a transport equation along the channel direction,

\[
-\frac{\hbar^2}{2m_x} \frac{\partial^2 \phi_m(x)}{\partial x^2} + E_m(x)\phi_m(x) = E\phi_m(x), \quad x \in (-\infty, +\infty).
\]

(6)

In the contacts, the electrons are assumed to be in an equilibrium state; thus both the potential \( V(x, z) \) and the subband energy function \( E_m(x) \) can be taken as constant there. As a result, \( E_m(x) \) can be assumed of the following form

\[
E_m(x) = \begin{cases} 
\nu_1, & -\infty < x \leq X_1, \\
\nu(x), & X_1 < x < X_2, \\
\nu_2, & X_2 \leq x < +\infty,
\end{cases}
\]

(7)

where \( [X_1, X_2] \) is identified as the device active region.

2.2. Density distribution and current formula

Due to the large scale in the \( y \) direction, the complete wave function \( \Phi(x, y, z) \) and the total energy \( E_{tot} \) are assumed to take the following forms with plane waves in the \( y \) direction,

\[
\Phi(x, y, z) = \Psi(x, z)e^{ik_y y}, \quad E_{tot} = \frac{\hbar^2 k_y^2}{2m_y} + E_z.
\]

According to the equilibrium Fermi–Dirac statistics, the density function is given by

\[
\rho(x, y, z) = \sum_{k_y} \sum_x F_j \left( \frac{\hbar^2 k_y^2}{2m_y} + E_z - \mu \right) \Psi_{-1/2}(x, y, z)\Psi_{1/2}(x, y, z),
\]

(8)

where \( \mu \) is the Fermi level and \( F_j \) is the Fermi function

\[
F_j(E - \mu) = \frac{1}{1 + e^{E - \mu}},
\]

with the Boltzmann constant \( k_B \) and the temperature \( T \). Summing over the wave-number \( k_y \) yields

\[
\rho(x, z) = \sum_x \sqrt{\frac{2m_y k_BT}{\pi \hbar^2}} F_{-1/2}(\mu - E_z)\Psi_{-1/2}(x, z)\Psi_{1/2}(x, z),
\]

(9)

where

\[
F_{\mu}(x) = \int_0^\infty t^{\mu} \frac{1}{1 + e^{(t-\mu)}} \, dt.
\]

(10)

Within the mode space method, the wave function \( \Psi_{\pm}(x, z) \) is assumed of a separable form

\[
\Psi_{\pm}(x, z) = \sum_m \Theta_m(x, z)\phi_m(x, E_z).
\]

As a result, the density function becomes

\[
\rho(x, z) = \sum_m \sum_x \bar{\rho}_m(x, E_z)|\Theta_m(x, z)|^2,
\]

(11)
where \( \hat{\rho}_m(x, E_z) \) is the density in the transport direction \( x \) for the \( m \)th mode,
\[
\hat{\rho}_m(x, E_z) = \sqrt{\frac{2m_k k_T}{\pi \hbar^2}} F_{-1/2}(\mu - E_z) |\phi_m(x, E_z)|^2.
\] (12)

Now, by replacing the discrete summation over \( x \) by a continuous integration and using the dispersion relation \( E_z = E_z(k_s) \),
\[
\sum_x \approx \frac{1}{2\pi} \int dk_s = \frac{1}{2\pi} \int \frac{dk_s}{dE_z} dE_z,
\] (13)

the density in the \( m \)th mode can be converted into an integration with respect to \( E_z \),
\[
\rho_m(x) = \sum_x \rho_m(x, E_z) \approx \int_0^{\infty} \rho_m(x, E_z) dE_z.
\] (14)

where
\[
\rho_m(x, E_z) = \sqrt{\frac{m_k k_T}{2\pi \hbar^2}} F_{-1/2}(\mu - E_z) |\phi_m(x, E_z)|^2 \frac{dk_s}{dE_z}.
\] (15)

Thus,
\[
\rho(x, z) = \sum_m \rho_m(x) |\Theta_m(x, z)|^2.
\] (16)

To compute the current transport through a device, we assume that free electrons at various energies are injected into the contacts and, based on the Landauer theory [9], the transport current can be computed through the reflection and transmission properties of the electron through the device. Mathematically, we assume the following set-up similar to an experimental setting. A unit amplitude plane wave is injected from the left (source) contact, and some portion of the electrons reflects from the device and some transmits through and exits the right (drain) contact without reflection [7].

Let \( k_1 = \sqrt{\frac{(2m_k k_T)^{-1}}{\pi \hbar^2}} \) and \( k_2 = \sqrt{\frac{(2m_k k_T)^{-1}}{\pi \hbar^2}} \). The wave functions in the left and the right contacts are of the form
\[
\begin{align*}
\phi_m(x) &= 1 e^{ik_1 x} + r_m e^{-ik_1 x}, \quad x \leq 0, \\
\phi_m(x) &= t_m e^{ik_2 x}, \quad x \geq L,
\end{align*}
\] (17)

where \( r_m \) and \( t_m \) are the reflection and the transmission coefficients for the source injection, and \( L \) is the length of the device. Here, we set \( X_1 = 0 \), and \( L = X_2 - X_1 \). The wave functions injected from the right drain contact have a similar form to the right of the device active region.

The probability flux in the \( m \)th mode is defined as
\[
j = \frac{\hbar}{2im_0} \left( \phi^*_m(x) \frac{\partial \phi_m(x)}{\partial x} - \phi_m(x) \frac{\partial \phi^*_m(x)}{\partial x} \right).
\]

It is easy to calculate the current transmission coefficient of free electrons from the source contact to the drain contact. For the \( m \)th mode, it is given by
\[
T_{m}^{d-s}(E_z) = \frac{\text{trans}}{\text{inc}} = 1 - |r_m|^2 = 1 - |\phi_m(0)|^2.
\] (18)

Therefore, the current due to the electrons with the energy \( E_z \) injected into the \( m \)th mode from the source contact is [9]
\[
l_m^{(\text{in})}(E_z) = e \sum_{h_y} T_{m}^{d-s}(E_z) F_j \left( \frac{\hbar^2 k_y^2}{2m_y} + E_z - \mu \right) v_x(E_z) = e \sqrt{\frac{2m_k k_T}{\pi \hbar^2}} F_{-1/2}(\mu - E_z) T_{m}^{d-s}(E_z) v_x(E_z).
\] (19)

According to the dispersion relation of the free electrons \( E_z = \frac{\hbar^2 k^2}{2m_k} \), \( v_x(E_z) = \frac{\hbar k}{m_k} \), and the procedure in (13), the total inflow current is
\[
l_m^{(\text{in})} = e \int_0^{\infty} \sqrt{\frac{m_k k_T}{2\pi \hbar^2}} F_{-1/2}(\mu - E_z) T_{m}^{d-s}(E_z) dE_z,
\] (20)

while the outflow current \( l_m^{(\text{out})} \) due to the electrons injected from the drain contact is analogous with (20). Finally, the total current for the \( m \)th mode is
\[
l_m = l_m^{(\text{in})} - l_m^{(\text{out})},
\] (21)

and
\[
l_m = \int_0^{\infty} I_m(E_z) dE_z,
\] (22)
where
\[
I_m(E_\varepsilon) = I_m^{\text{in}}(E_\varepsilon) - I_m^{\text{out}}(E_\varepsilon) = \frac{e}{\hbar} \sqrt{\frac{m_e k_B T}{2 \pi^3}} \left[ F_{-1/2}(\mu_i - E_\varepsilon) - F_{-1/2}(\mu_d - E_\varepsilon) \right] T_{m}^{-d}(E_\varepsilon). \tag{23}
\]

3. Computing transport current with the NEGF

3.1. Boundary conditions of the Green’s function

The NEGF method can be used to solve the infinite domain problem (6) by defining the Green’s function \( G(x,x') \) as the solution to the following problem
\[
\left( E - E_m(x) + \frac{\hbar^2}{2 m_e} \frac{\partial^2}{\partial x'^2} \right) G(x,x') = \delta(x-x'), \quad x,x' \in (-\infty, +\infty). \tag{24}
\]
Boundary conditions suitable for the FDM (Finite Difference Method) and the FEM (Finite Element Method) have been derived with a unified treatment for the NEGF [20]. If the FDM is used to solve the NEGF equation, the left boundary condition is
\[
G(x'_c, x') = e^{-i k_c (x'_c - x_1)} G(X_1, x'), \quad x'_c \in (-\infty, X_1), \quad x' \in [X_1, X_2], \tag{25}
\]
which can be used to compute the value of \( G(x'_c, x') \) for \( x'_c \) outside the FDM computational domain \([X_1, X_2]\). Meanwhile, the right boundary condition is
\[
G(x'_c, x') = e^{i k_c (x'_c - x_1)} G(X_2, x'), \quad x'_c \in (X_2, +\infty), \quad x' \in [X_1, X_2]. \tag{26}
\]

The boundary conditions for the Green’s function are designed to observe the causality of a retarded Green’s function such that the electron waves leaving the quantum device region into the contacts will not experience reflections at the device boundaries. The derivation of (25) and (26) is based on the Sommerfeld outgoing radiation condition imposed on the Green’s function \( G(x,x') \), i.e., \( \left( \frac{d}{dx} \pm ik_c \right) G(x,x') \rightarrow 0 \), as \( x \rightarrow \mp \infty \), which implies that the electron wave due to a local Dirac excitation will propagate to the infinite without reflection. The details can be found in Eqs. (35)–(37) in [20].

3.2. Current formula by the NEGF

In the framework of the NEGF, the density matrix of the electrons injected with the energy \( E_\varepsilon \) into the \( m \)th subband is [2,6]
\[
\rho_m(E_\varepsilon) = \frac{1}{a} \sqrt{\frac{m_e k_B T}{2 \pi^3 \hbar^2}} [ F_{-1/2}(\mu_i - E_\varepsilon) A_s + F_{-1/2}(\mu_d - E_\varepsilon) A_d ], \tag{27}
\]
where the diagonal element gives just the density \( \rho_m(x, E_\varepsilon) \) defined in (15), and \( A_s, A_d \) are the spectral functions related to the Green function
\[
A_s = G \Gamma_s G^+, \quad A_d = G \Gamma_d G^+, \tag{28}
\]
with
\[
G = [EI - H - \Sigma_s - \Sigma_d]^{-1}, \quad \Gamma_s = i \left( \Sigma_s - \Sigma_s^+ \right), \quad \Gamma_d = i \left( \Sigma_d - \Sigma_d^+ \right). \tag{29}
\]
Here, \( \Sigma_s, \Sigma_d \) are defined as the self-energies and describe the effect of the source and drain contacts to the device. In fact, the self-energy is related to the boundary condition of Green’s function (25) and (26) [20]. \( \Gamma_s \) and \( \Gamma_d \) are related to the velocities at the injection and exit points.

We will use a second order central finite difference scheme to solve (24) on \([X_1, X_2]\). Let \( a = \frac{x_2 - x_1}{N} \), \( x_0 = X_1 \), \( x_i = X_1 + ia \), \( x_N = X_2 \). Then, the corresponding matrix \( EI - H \) is given by
\[
EI - H = \begin{pmatrix}
A_0 & t_x & 0 & \cdots & \cdots & \cdots \\
t_x & A_1 & t_x & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\end{pmatrix},
\]
in which \( t_x = \frac{\hbar^2}{2 m_e a^2} \), \( A_i = E - 2t_x - E_m(x_i) \). And, the self-energy \( \Sigma_{s,d} \) is expressed as [20]
Algorithm 1. A Mode Space Method with the NEGF for the channel transport

Given an error tolerance $\epsilon > 0$.

1. Guess an initial potential function $V(x, z)$;
2. Solve the eigenvalue problem (5) at each slice $x = x_s$. Calculate eigenvalues $E_m(x_s)$ form the subband $E_m(x)$ in the $x$ direction. At least three subbands will be calculated;
3. Solve the transport equation at each subband $E_m(x)$. With the NEGF method (24)–(26) by a second order central finite difference method on $[x_1, x_2]$ to calculate the density matrix $\rho_m(E)$ via Eq. (27). The diagonal element is the density matrix for the injected electrons with energy $E$;
4. Insert the density $\rho_m(E)$ into Eq. (34) and integrate over energy $E$ to obtain the electron density $\rho(x, z)$;
5. Solve the Poisson Eq. (1) with a Newton iteration method [20]. With the updated potential $V(x, z)$, repeat Step 2 to Step 5 until the potential distribution $V(x, z)$ is convergent within the given error tolerance $\epsilon$;
6. Compute the current with the convergent potential according to (33), (23) and (30).

Remark 1. In our computations, it is effective to set zeros as the initial guess on potential for small gate bias, especially for $V_g = 0$. To investigate the $I$–$V$ curve, we need to obtain a group of values of voltages and currents $(V_g, I_j)$, $j = 0, 1, 2, \ldots$. We set $V_g^0 = j \times 0.1$ V in our computation. The initial guess value on potential for $V_g^j$. $j > 0$ is the convergent potential for $V_g^{j-1}$. This strategy is found to be effective in our simulations.
Remark 2. Because of the symmetry of the eigenvalue problem (5) and that only a few small eigenvalues are needed, bisection method is applied to obtain the eigenvalues in Step (2).

Remark 3. For integration with respect to energy $E$ for the computation of the density and the current in Step (4) and Step (6), an adaptive Romberg’s method is applied to ensure the density and the current convergent within a given error tolerance.

4. Computing transport current with the Wigner equation

4.1. Wigner equations

The Wigner function is defined through a Fourier transform for the density matrix, which is, for the $m$th subband, 

$$\rho_m(x, x') = \sum_{x} \frac{2m_{k_0}}{\pi h^2} F_{-1/2}(\mu - E_z) \phi_m(x, E_z) \phi_m^*(x', E_z),$$

(35)

where the diagonal element $\rho_m(x, x)$ is the density function for the $m$th subband. With a Weyl transform

$$x = R + \frac{r}{2}, \quad x' = R - \frac{r}{2},$$

the Wigner distribution function for the $m$th mode is defined by

$$f_w^m(R, q) = \int_{-\infty}^{+\infty} \rho_m \left( R + \frac{r}{2}, R - \frac{r}{2} \right) e^{-iqr} dr.$$

(36)

Meanwhile, the Wigner function corresponding to a wave function $\phi_m(x, E_z)$ is

$$f_w^{m,z}(R, q) = \int_{-\infty}^{+\infty} \phi_m \left( R + \frac{r}{2}, E_z \right) \phi_m^* \left( R - \frac{r}{2}, E_z \right) e^{-iqr} dr.$$

(37)

The density $\rho_m(x)$ is related to the Wigner function $f_w^m(x, q)$ by

$$\rho_m(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f_w^m(x, q) dq.$$

(38)

To derive the equation for the Wigner function $f_w^m(x, q)$ and the appropriate boundary conditions for an open system, we apply the Weyl transform and Fourier transform with respect to $r$ to both sides of Eq. (35) to get

$$f_w^m(x, q) = \sum_{x} \frac{2m_{k_0}}{\pi h^2} F_{-1/2}(\mu - E_z) f_w^{m,z}(x, q).$$

(39)

The equation satisfied by the Wigner function $f_w^{m,z}(x, q)$ can be derived from (6) with the Weyl and the Fourier transforms [3–5,13]

$$\frac{h_{q}}{m_{k}} \frac{\partial}{\partial x} f_w^{m,z}(x, q) + \frac{1}{\pi h} \int_{-\infty}^{+\infty} V_w(x, q - q') f_w^{m,z}(x, q') dq' = 0,$$

(40)

where the Wigner potential

$$V_w(x, q - q') = \int_{0}^{+\infty} \sin((q - q')r) [E_m(x + \frac{r}{2}) - E_m(x - \frac{r}{2})] dr.$$

(41)

From Eq. (39), the Wigner function $f_w^m(x, q)$ will satisfy

$$\frac{h_{q}}{m_{k}} \frac{\partial}{\partial x} f_w^m(x, q) + \frac{1}{\pi h} \int_{-\infty}^{+\infty} V_w(x, q - q') f_w^m(x, q') dq' = 0.$$  

(42)

4.2. Boundary conditions of the Wigner function

The Wigner Eq. (42) contains a first order derivative in $x$; therefore, only one boundary condition is required at the boundary of a finite domain. Because the potential $E_m(x)$ in the contacts of both sides is assumed to be uniform, the solutions in the semi-infinite contacts are plane waves of the form $\phi_m(x)$ as in Eq. (17) in the left semi-infinite domain, for instance. According to the definition (36) of the Wigner distribution, for illustrative argument, as a rough approximation deep inside the left contact, we have

$$f_w^{m,z}(x, q) = \int_{-\infty}^{+\infty} \phi_m \left( x + \frac{r}{2} \right) \phi_m^* \left( x - \frac{r}{2} \right) e^{-iqr} dr = \delta(k_1 - q) + |r_m|^2 \delta(k_1 + q) - 2r_m i \sin(k_1 x) \delta(q), \quad (k_1 > 0, \; x < 0).$$

(43)
The inflow boundary condition proposed in [4] at the left boundary specifies \( f_m^{x-}(0, q) \) when \( q > 0 \), and at the right boundary \( f_m^{x-}(L, q) \) when \( q < 0 \). Eq. (43) shows that only the injection wave contributes to the left boundary condition and the reflection wave has no effect on it (due to the fact that \( \delta(k_1 + q) = 0 \), \( \delta(q) = 0 \) for \( k_1 > 0 \), \( q > 0 \)). The right boundary condition assumes a unit wave injected from the right contact. Next, for open systems, free electrons are supposed to be injected from infinity, so the energy \( E_z \) of the free electrons injected from the left contact has the form,

\[
E_x = \frac{\hbar^2 k_1^2}{2m_x} + v_1.
\]

Therefore, according to Eqs. (39) and (43), the left boundary condition for the Wigner function should be

\[
f_w^m(0, q) = \sqrt{\frac{2m_x k_0 F}{\pi \hbar^2}} F_{-1/2} \left( \mu - k_0^2 q^2 \frac{2m_x}{2m_x} - v_1 \right), \quad q > 0.
\]

On the other hand, for a unit amplitude wave injected from the right, we can obtain the right boundary condition similarly

\[
f_w^m(L, q) = \sqrt{\frac{2m_x k_0 F}{\pi \hbar^2}} F_{-1/2} \left( \mu - k_0^2 q^2 \frac{2m_x}{2m_x} - v_2 \right), \quad q < 0.
\]

If the Fermi energy levels of the two contacts are different, electrons injected from the left contact are in equilibrium with the left contact and those from the right contact are in equilibrium with the right contact. Thus, the boundary conditions for a non-equilibrium state is

\[
\begin{align*}
&f_w^m(0, q) = \sqrt{\frac{2m_x k_0 F}{\pi \hbar^2}} F_{-1/2} \left( \mu - k_0^2 q^2 \frac{2m_x}{2m_x} - E_m(x_1) \right), \quad q > 0, \\
&f_w^m(L, q) = \sqrt{\frac{2m_x k_0 F}{\pi \hbar^2}} F_{-1/2} \left( \mu - k_0^2 q^2 \frac{2m_x}{2m_x} - E_m(x_2) \right), \quad q < 0,
\end{align*}
\]

where \( \mu_r, \mu_d \) are the Fermi energy of the source contact and the drain contact, respectively.

4.3. Current formula with Wigner distributions

After calculating \( f_m(x, q) \) by Eqs. (42) and (46), the density \( \rho_m(x) \) can be obtained by Eq. (38). Then, summing over all sub-bands, we obtain the total density

\[
\rho(x, z) = \frac{1}{B} \sum_m \rho_m(x) |\Theta_m(x, z)|^2.
\]

The total current density is

\[
j(x) = \frac{1}{2\pi} \sum_m I_m(x),
\]

where

\[
I_m(x) = e \int_{-\infty}^{+\infty} \frac{\hbar q}{m_x} f_w^m(x, q) dq.
\]

For static problems, the current should be independent of \( x \). According to Eqs. (39) and (43), the current due to the source injection is

\[
I_m(x, k_1) = e \sqrt{\frac{2m_x k_0 F}{\pi \hbar^2}} F_{-1/2}(\mu - E_x)(1 - |r_m(E_x)|^2) \frac{\hbar k_1}{m_x}.
\]

Recalling that for the Schrödinger wave description of scattering (18), we have

\[
T_m^{s-d} = 1 - |r_m(E_x)|^2, \quad v_s(E_x) = \frac{\hbar k_1}{m_x}.
\]

So \( I_m(x, k_1) \) is exactly \( I_m(E_x) \) in Eq. (23). Therefore, the density function and the current formula described by the Wigner function are equivalent to those in the NEGF.

4.4. Numerical scheme for the Wigner equation

The Wigner equation is solved in the \((x, q)\) space. Let \( h_x \) be the mesh size of the \( x \) space, and \( N + 1 \) be the number of mesh points,
\[ h_x = \frac{L}{N}, \quad x_i = \frac{L}{2} + ih_x, \quad i = 0, 1, 2, \ldots, N. \]

From Eq. (42), only the boundary condition in the \( x \)-space is required. A upwind difference method is used to approximate first order derivative in \( x \). We denote \( L_q \) the integration length in \( q \), and its value will be determined by the mass conservation property requirement below. \( N_q \) is the number of mesh points for the variable \( q \). In order to avoid \( q = 0 \) which would lead to a zero element in the diagonal of the matrix, we set the mesh points as

\[ h_q = \frac{L_q}{N_q}, \quad q_j = \frac{L_q}{2} - \left( j + \frac{1}{2} \right) h_q, \quad j = 0, 1, 2, \ldots, N_q - 1. \]

Discretizing the Wigner function at \( (x_i, q_j) \) yields

\[
\begin{align*}
\frac{h_q}{m} f_w(x_i) - f_w(x_{i-1}) &= \frac{1}{\pi h_x} \sum_{j=0}^{N_q-1} V_w(x_i, q_j - q_j') f_w(x_i, q_j') = 0, & q_j > 0, \\
\frac{h_q}{m} f_w(x_{i-1}) - f_w(x_i) &= \frac{1}{\pi h_x} \sum_{j=0}^{N_q-1} V_w(x_i, q_j - q_j') f_w(x_i, q_j') = 0, & q_j < 0,
\end{align*}
\]

for \( i = 1, 2, 3, \ldots, N - 1 \), and \( j = 0, 1, 2, \ldots, N_q - 1 \). Here \( V_w(x_i, q_j - q_j') \) is calculated by a numerical integration of Eq. (41). In fact, if the phase-breaking process is so effective inside the contacts to destroy completely the coherence, the correlation between any point inside the active region and a point inside the contact regions can be set to zero. So the integration interval for \( r \) in Eq. (41) is limited to \( \frac{1}{2} - |q| \) [22].

\[
V_w(x, q - q') = \int_0^{\frac{\pi}{2}} \sin((q - q')r) \left( E_m(x + \frac{r}{2}) - E_m(x - \frac{r}{2}) \right) dr.
\]

In order to maintain the carrier conservation for each subband (assuming no charge scattering between different subbands), the following integration over the \( q \)-space should be zero, which amounts to saying that the net scattering of electrons between \( q \) and \( q' \) should be zero for conservation,

\[
\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} dq \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V_w(x, q - q') f_m^w(x, q') dq' = 0.
\]

Here, we have zeroed-out the distribution function \( f_m^w(x, q) = 0 \) if \( |q| > \frac{L_q}{2} \), and effectively, Eq. (42) is only needed for \( |q| < \frac{L_q}{2} \) [16]. Now, as \( V_w(x, q - q') \) is calculated by a numerical quadrature in the \( r \) variable (see (57) below), Eq. (53) implies that

\[
\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin((q - q')r) dq = 0.
\]

Therefore, we have

\[
\cos \left( \frac{L_q}{2} - q' \right) r - \cos \left( \frac{L_q}{2} - q \right) r - L_q r = 0.
\]

Consequently, we conclude that \( L_q r \) should be a multiple of \( 2\pi \). Let \( h_r \) denote the mesh size for the variable \( r \) and \( r_k = kh_r \). Thus, \( L_q \) and \( h_r \) should satisfy the following relation in order to ensure Eq. (55)

\[
L_q h_r = 2\pi, \quad L_q = \frac{2\pi}{h_r},
\]

which defines the integration length of \( L_q \). Numerical integration is now applied to (41). For convenience, we set \( h_r = 2h_x \). Then, we have

\[
V_w(x_i, q_j - q_j') = h_r \sum_{k=0}^{N_q} \sin \left( \frac{2k(j - j')\pi}{N_q} \right) \left( E_m(x_i, k) - E_m(x_i, k) \right).
\]

In order to improve the computation accuracy, \( h_r \) can be reduced to \( h_r = h_x \) or \( h_r = \frac{h_x}{2} \). If \( E_m(x + \frac{r}{2}) \) is not at a mesh point, a linear interpolation is used to obtain it. The current density is computed as,

\[
j \left( x + \frac{h_x}{2} \right) = \sum_m h_q \sum_{q_j < 0} \frac{h_q}{m_x} f_m^w(x + h_x, q_j) + \sum_{q_j > 0} \frac{h_q}{m_x} f_m^w(x, q_j).
\]

This definition ensures that the current density calculated for the steady-state solution is independent of \( x \) [4].

Finally, the mode space method with the Wigner transport equation gives the following algorithm.

**Algorithm 2.** A Mode Space Method with the Wigner equation for the channel transport
Given an error tolerance \( \epsilon > 0 \).

1. Guess an initial potential function \( V(x, z) \);
2. Solve the eigenvalue problem (5) at each slice \( x = x_i \). Calculate eigenvalues \( E_m(x_i) \) form the subband \( E_m(x) \) in the \( x \) direction. Again, at least three subbands are calculated;
3. Solve the transport equation at each subband \( E_m(x) \) with the Wigner equation method (42) to calculate the density \( \rho_m(x) \) via Eq. (38);
4. Insert the density \( \rho_m(x) \) into Eq. (47) to obtain the electron density \( \rho(x, z) \);
5. Solve the Poisson Eq. (1) with a Newton iteration method [20]. With updated potential \( V(x, z) \), repeat Step 2 to Step 5 until the potential distribution \( V(x, z) \) is convergent within the given error tolerance \( \epsilon \);
6. Solve the eigenvalue problem (5) and the Wigner Eq. (42) with the convergent potential \( V(x, z) \), and calculate the current by (58).

5. Simulation results of a nano-scale DG-MOSFET

We simulate a symmetric DG-MOSFET with a 9 nm gate length, a 3 nm thickness of silicon and a 1 nm thickness of oxides. The doping density in the source and drain contacts is \( 2 \times 10^{20} \text{ cm}^{-3} \), damped exponentially into the channel with intrinsic doping \( 1 \times 10^{16} \text{ cm}^{-3} \). The mesh size is set to \( a = h_x = 0.3 \text{ nm} \), and \( b = 0.1 \text{ nm} \). In the Wigner function calculation, \( h_r \) determines the integration region \( L_q \) given in Eq. (56). It is more accurate to use smaller \( h_r \), though at a larger cost. Fig. 2 shows that the accuracy is improved when \( h_r \) is refined, but the computation cost increases greatly. By balancing the accuracy and computational cost, we set \( h_r = 2 h_x \).

In both the NEGF and the Wigner equation methods, a numerical mesh is used to solve the differential equations. The numerical results presented here have achieved mesh convergence for both methods. In the NEGF method, complex linear systems have to be solved for each energy value \( E \) while we have to solve a differential–integral equation for the Wigner distribution in the momentum and position space. For convenience, a uniform mesh is adopted in the computation of the Wigner distribution and the wave-number mesh \( h_q \) has to be small enough to ensure the accuracy. While in the NEGF, an adaptive Romberg’s method is applied to integrate the energy \( E \) and to improve the computational efficiency. In fact, the NEGF method is more efficient for ballistical transport simulations of 2D DG-MOSFET than the Wigner method. However, the Wigner equation can easily include the scattering effect by adding the scattering term similar to that of the Boltzmann equation.

Figs. 3 and 4 are the density distribution and the potential distribution with the NEGF method (denoted by NEGF) and the Wigner function method (denoted by Wigner). Obviously, these two methods produce the same density and the potential distributions under given gate and source–drain voltages. In Fig. 5, plot A is the first subband profile by the two methods, and it is easy to see that the subbands of two methods almost overlap. Plot B is the Wigner function, which is sharp and has negative values at the potential barrier, indicating the potential barrier scattering effect on the electrons.

From the density function, the potential distribution and the subband energy, we can not distinguish the difference between the NEGF method and the Wigner distribution method. As an \( I-V \) (voltage and current) curve can be measured experimentally, so it is more significant to investigate the \( I-V \) curve of the two methods. \( I_{ds} \) is the current and \( V_{ds} \) is the voltage.

![Fig. 2. Error of \( \rho(x) \) for different \( h_r \) at \( V_g = 0 \text{ V} \) and \( V_{ds} = 0 \text{ V} \) with the 1D Wigner transport equation.](image-url)
between the source and the drain. In our simulations, the source contact is set to be the potential benchmark. \( V_g \) is the gate voltage. Fig. 6 shows the \( I-V \) curves of \( I_{ds} \) vs \( V_{ds} \) with different gate voltages (the left one for \( V_g = 0.538 \) V, the right one for \( V_g = 0.738 \) V). The simulation results indicate that the \( I-V \) curves calculated by the Wigner function method are slightly
higher than those by the NEGF method. From Fig. 6, however, we can also see that the $I-V$ curves with these two methods approach each other as $V_g$ increases.

Fig. 7 shows the $I-V$ curves illustrating how the source–drain current $I_{ds}$ changes with the gate voltage $V_g$ for $V_{ds} = 0.05$ V (the left one) and $V_{ds} = 0.4$ V (the right one). At low source–drain voltages, the $I-V$ curves of the two methods almost overlap. With increasing source–drain voltages, the current calculated by the NEGF method is again lower than that by the Wigner function method.

We also investigate the effect of the thickness of the silicon body and the oxide layers to the $I-V$ curve in Fig. 8, which shows the $I-V$ curves getting higher when the silicon layer becomes thicker. If we double the silicon layer and the oxide layer simultaneously, the $I-V$ curves are lower than the original ones when $V_{ds}$ is less than 6 V and higher than original ones with $V_{ds}$ greater than 6 V.

Although the current calculated by the Wigner equation is a little higher than that calculated by the NEGF method, the $I-V$ curves of the two methods have the same trend.
6. Conclusion

In this paper, we have shown two different ways quantum device boundaries are treated in the NEGF and the Wigner equation methods for the transport along the channel direction in a nano-scale DG-MOSFET. In the NEGF method, an outgoing boundary condition is imposed to ensure the casualty of the retarded Green's function such that electron waves leaving the device active region into the contacts will not suffer reflections. On the other hand, the Wigner equation is equipped with an inflow boundary condition such that free electrons from the contacts can enter the quantum device active region without reflections at the boundaries. Both methods are integrated into the mode space framework. Due to the difference of the boundary treatments, the transport current through the DG-MOSFET shows a higher through current for the Wigner equation than that computed by the NEGF. Also, we have found that the thickness of the silicon and the oxide affects much the $I$–$V$ characteristics of the device.

We conclude this paper by discussing the possible reasons why the two methods with their boundary treatments show some differences in the $I$–$V$ characteristics. In the treatment of the boundary condition for the NEGF, the boundary conditions (25) and (26) rely on the ideal non-reflection of electrons emitting from the device into the equal potential contacts and the assumption that the incident waves into the device are given in the Fermi–Dirac distributions. Both assumptions should be further investigated, say, by including more contact regions into the simulated device region. On the other hand, the inflow boundary condition for the Wigner equation is a local condition, and it has been pointed out in [24] that in fact the Wigner potential depends globally on its previous time in the phase space through the integral operator. This fact is a current active research topic. Moreover, the justification in (43) for giving the inflow boundary condition only applies far away from the device and deep into the contact region. The plane wave assumption is only valid to one side of the device while the Wigner distribution involves the correlation of the density operator over the whole one-dimensional space. As a result, pre-assigning the Fermi–Dirac distribution to the Wigner distribution at the device boundaries, which may be related to the higher current, requires further investigations to quantify its effect on the transport current. All these issues will be examined in future research to understand the source for the effects of the boundary conditions on the transport current.

Acknowledgments

The authors thank the support of the US Army Research Office (Grant No: W911NF-07-1-0492). H.Y. Jiang is also partially supported by National Science Foundation of China for Tianyuan Young Scholars (Grant No. 10726025) and by the Duke post-doctoral fellowship from Charlotte Research Institute. W. Cai also acknowledges the support of NSFC (Grant No. 10828101).

References


