Chapter 2
Quantum and Coulomb Effects in Nano Devices

Dragica Vasileska, Hasanur Rahman Khan, Shaikh Shahid Ahmed, Gokula Kannan, and Christian Ringhofer

Abstract In state of the art devices, it is well known that quantum and Coulomb effects play significant role on the device operation. In this book chapter we demonstrate that a novel effective potential approach in conjunction with a Monte Carlo device simulation scheme can accurately capture the quantum-mechanical size quantization effects. Inclusion of tunneling within semi-classical simulation schemes is discussed in details. We also demonstrate, via proper treatment of the short-range Coulomb interactions, that there will be significant variation in device design parameters for devices fabricated on the same chip due to the presence of unintentional dopant atoms at random locations within the channel of alternative technology devices.

Keywords Nanoscale devices · Quantum confinement · SCHRED · Random dopants

1 Introduction

As semiconductor devices are being scaled into nanometer dimensions (Fig. 2.1), significant number of effects start to become important and they can be classified into quantum and classical reliability effects. In general, there are three manifestations of quantum effects in nanodevices: (1) quantum-mechanical size quantization, (2) tunneling and (3) quantum interference. Quantum-mechanical size quantization effects and gate leakage can be easily incorporated into classical simulators, but quantum interference effects require fully quantum-mechanical treatment. In this book chapter we focus on the inclusion of quantum-mechanical size quantization and tunneling effects into particle-based device simulators. Several separate book chapters in this book are devoted to quantum transport. In addition

D. Vasileska (✉)
School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, USA
e-mail: vasileska@asu.edu

D. Vasileska and S.M. Goodnick (eds.), *Nano-Electronic Devices: Semiclassical and Quantum Transport Modeling*, DOI 10.1007/978-1-4419-8840-9_2,

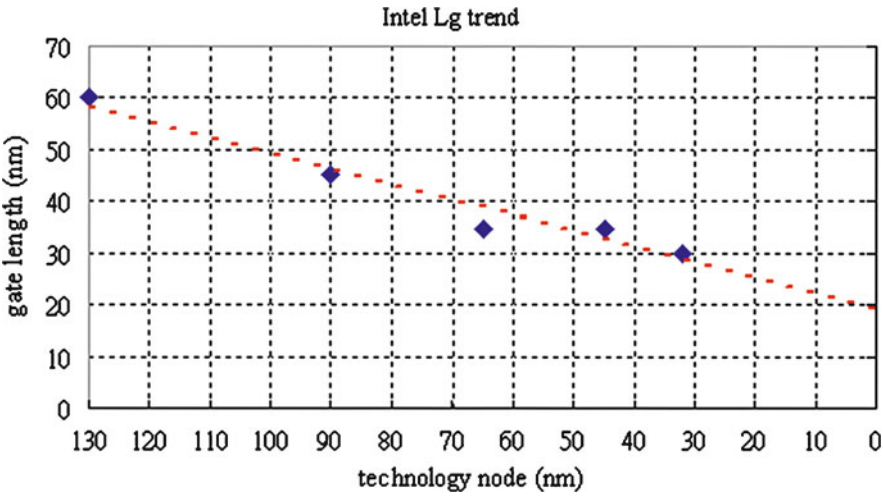


Fig. 2.1 Intel trend in transistor channel length scaling

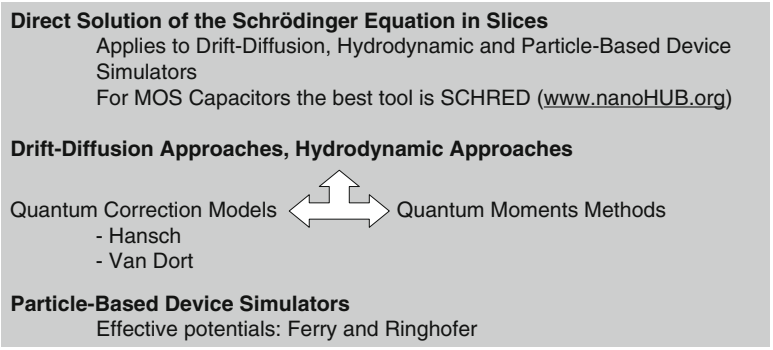


Fig. 2.2 Inclusion of Quantum Mechanical Space/Size Quantization effects in classical device simulators

to this, in this book chapter we also address in detail the issue of transistor reliability due to random dopant effects or due to unintentional dopants in alternative technology devices.

The inclusion of quantum-mechanical size quantization effects in drift-diffusion, hydrodynamic and particle-based device simulators is schematically illustrated in Fig. 2.2 and explained in more detail later in the text.

Quantum correction models try to incorporate quantum-mechanical description of carrier behavior via modification of certain device parameters within the standard drift-diffusion or hydrodynamic model. For example, the Hansch model [1] modifies the effective density of states function using,

$$N_C^* = N_C[1 - \exp(-z/\mathrm{LAMBDA})]^2 \tag{2.1}$$

where LAMBDA is a parameter.

On the other hand, the very popular Van Dort model [2] modifies the intrinsic carrier concentration by taking into account the effective band-gap increase due to quantum-mechanical size quantization effects. Namely, the surface potential is modified according to:

$$\psi_s^{QM} = \psi_s^{CONV} + \Delta\varepsilon/q + E_n \Delta z,\ \Delta z = \langle z^{QM} \rangle - \langle z^{CONV} \rangle \tag{2.2}$$

The second term on the RHS of the above expression accounts for the band-gap widening effect because of the upward shift of the lowest allowed state. The third term accounts for the larger displacement of the carriers from the interface and the extra band-bending needed for given population that is expressed with

$$qE_n \Delta z \approx \frac{4}{9}\Delta\varepsilon \tag{2.3}$$

The energy shift that appears in the above equation is calculated using the variational approach of Fang and Howard [3]. With these modifications, one arrives at the following expression for the effective band-gap

$$E_g^{QM} = E_g^{CONV} + \frac{13}{9}\Delta\varepsilon,\ \Delta\varepsilon \approx \beta \left(\frac{\varepsilon_{Si}}{4qk_BT} \right)^{1/3} E_\perp^{2/3} \tag{2.4}$$

where β is a parameter. The modification in the effective bandgap leads to modification of the intrinsic carrier concentration

$$\begin{aligned} n_i^{QM} &= n_i^{CONV} \exp\left[(E_g^{QM} - E_g^{CONV})/2k_BT \right] \\ n_i &= n_i^{CONV}[1 - F(y)] + F(y) n_i^{QM} \end{aligned} \tag{2.5}$$

where the function $F(y)$ defined with

$$F(y) = 2\exp(-a^2)/\left[1 + \exp(-2a^2)\right],\quad a = y/y_{ref} \tag{2.6}$$

enables a smooth transition between the intrinsic carrier density in the quantum region (towards the semiconductor-oxide interface) and the semiclassical region (towards the bulk portion of the device). The meaning of the various parameters that appear in the expressions of the Van Dort model is graphically represented in Fig. 2.3 below.

The quantum moment methods for inclusion of size quantization effects into drift-diffusion and hydrodynamic simulators are discussed in Sect. 2.1 below. SCHRED First and Second Generation are discussed in Sect. 2.2. SCHRED First Generation (or SCHRED V1.0) is a tool developed by Prof. Vasileska from Arizona State University back in 1992 and it was further developed in 1998 and installed on PUNCH (in fact, SCHRED was the first tool installed on Purdue University Network Computational Hub). When the Network for Computational Nanotechnology (NCN) was formed, SCHRED V1.0 was immediately transferred on the

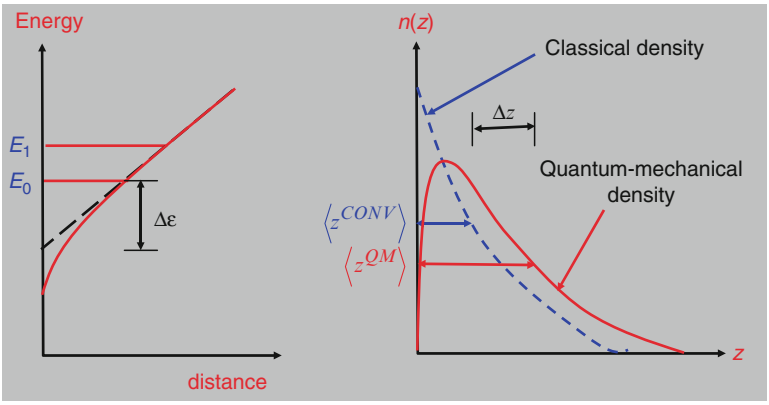


Fig. 2.3 Graphical description of the idea of the Van Dort model

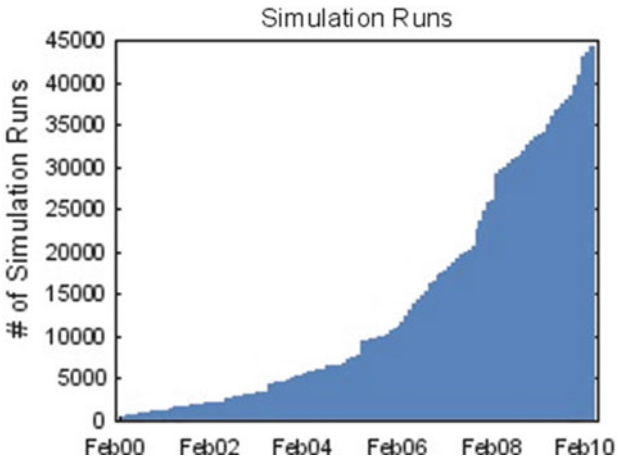


Fig. 2.4 SCHRED usage statistics

nanoHUB portal. In the meantime SCHRED V1.0 went through several revisions made by Prof. Vasileska and Dr. Zhibin Ren (Currently at IBM T. J. Watson), the most important being the introduction of quantization of holes using a heavy-hole and a light-hole band model and calculation of the tunneling current through the gate oxide. After being installed on PUNCH, and more so after its installment on the nanoHUB, SCHRED V1.0 gained enormous popularity. In fact, it was not only popular for educators to help teach students principles of operation of MOS capacitors, it was also heavily used in research work all around the world and is at the moment cited in 108 research papers (www.nanoHUB.org). The usage statistic of SCHRED v1.0 is depicted in Fig. 2.4 and its world-wide usage is illustrated in Fig. 2.5.

Fig. 2.5 SCHRED worldwide usage

The trend in transistors channel length scaling shown in Fig. 2.1 also requires oxide thickness reduction to improve the device transconductance and achieve better control of the charge in the channel with the gate. Since 1 nm oxide have shown to be very leaky, Intel in its 45 nm technology node already introduced high-k dielectrics, thus eliminating the gate leakage problem.

However, the gate leakage is still a big issue in Schottky transisors like MESFETs and HEMTs. The calculation of the gate leakage current in these structures can be accomplished by the use of either the WKB approximation or the transfer matrix approach. With regard to the injection between the Schottky gate and the device channel, it is best handled by using transmission probabilities, which are obtained as solutions of the Schrödinger equation along paths perpendicular to the semiconductor/metal interface. The potential along these paths is taken from the solution of the Poisson equation at each self-consistent step of the Monte Carlo procedure. The transmission probability is calculated using standard Airy function approach based on the 1 D Schrödinger equation on the propagating path. A transfer matrix approach is then applied, where the potential is interpolated linearly between the grid points on which the Poisson equation is solved in the Monte Carlo region. The unique solution is calculated with the application of the boundary conditions for the continuity of the wavefunction and its derivative at each grid point. The use of the Airy functions approach is better than the simple WKB approximation, because WKB model neglects quantum-mechanical reflections for the thermionic emission and is typically inaccurate for tunneling near the top of the potential barrier. Direct solution of the Schrödinger equation, as implemented via the Airy function formalism, also has the advantage of treating on an equal footing both thermionic emission and field-emission tunneling.

To compute the current injected by the metal contact, we calculate transmission coefficient as a ratio of the transmitted and incident probability current densities. At each iteration step, a table of transmission probabilities is generated for each mesh location along the contact interface. Then, the injected current density is obtained by integrating the product between carrier distribution and transmission probability. In its actual implementation within the Monte Carlo scheme, the transmission probability is evaluated separately for each particle and a random number technique is used to decide whether the particle is absorbed or not. Note that a similar version of the above-described approach has been successfully applied in simulations of Schottky barrier MOSFETs, as described in more detail in [4]. The WKB approximation and the transfer matrix approach that employs Airy function solutions for piecewise linear potential barrier are explained in Sect. 3 of this book chapter.

Yet another issue that we discuss in this book chapter in great details is transistor mismatch due to random number and random position of the impurity atoms in the active region of the device. These statistical fluctuations of the channel dopant number were predicted by Keyes [5] as a fundamental physical limitation of MOSFET down-scaling. Entering into the nanometer regime results in a decreasing number of channel impurities whose random distribution leads to significant fluctuations of the threshold voltage and off-state leakage current. These effects are likely to induce serious problems on the operation and performances of logical and analog circuits. It has been experimentally verified by Mizuno and co-workers [6] that threshold voltage fluctuations are mainly caused by random fluctuations of the number of dopant atoms and that other contributions such as fluctuations of the oxide thickness are comparably very small. It follows from these remarks that impurities cannot be considered anymore using the continuum doping model in advanced semiconductor device modeling but the precise location of each individual impurity within a full Coulomb interaction picture must be taken into account.

In the past, the effect of discrete dopant random distribution in MOSFET channel has been assessed by analytical or drift-diffusion (DD) approaches. The first DD study consisted in using a stochastically fluctuating dopant distribution obeying Poisson statistics [7]. 3D *atomistic* simulators have also been developed for studying threshold voltage fluctuations [8,9]. Even though the DD/HD methods are very useful because of their simplicity and fast computing times, it is not at all clear whether such macroscopic simulation schemes can be exploited into the atomistic regime. In fact, it is not at all clear how such discrete electrons and impurities are modeled in macroscopic device simulations due to the long-range nature of the Coulomb potential.

Three-dimensional (3D) Monte Carlo (MC) simulations should provide a more realistic transport description in ultra-short MOSFETs. The MC procedure gives an exact solution of the Boltzmann transport equation. Thus it correctly describes the non-stationary transport conditions. Even where the microscopic simulations such as the MC method are considered, the treatment of the electrons and impurities is not straightforward which is again due to the long-range nature of the Coulomb potential. The incorporation of the long-range Coulomb potential in the MC method has been a long-standing issue [10,11]. This problem is, in general, avoided by assuming

that the electrons and the impurities are always screened by the other carriers so that the long-range part of the Coulomb interaction is effectively suppressed. The complexity of the MC simulation increases as one takes into account more complicated screening processes by using the dynamical and wave-vector dependent dielectric function obtained from, for example, the random phase approximation. However, the screening is a very complicated many-body matter [12].

This situation is also complicated in the MC *device simulations* in which the BTE is self-consistently coupled with the Poisson equation [13]. The Coulomb potential due to electrons and impurities is then separated into the long-range and the short-range parts. The long-range part is taken into account by the solution of the Poisson equation, whereas the short-range part is usually included in the BTE through the scattering kernel. In other words, the Coulomb potential is separated into the long-range and short-range parts by the size of the mesh employed in the Poisson equation. However, the choice of the mesh size is not trivial. For example, the mesh cannot be arbitrarily small as the Coulomb potential would then be double-counted by the Poisson equation and the BTE. Since the long-range part of the Coulomb potential is responsible for the many-body effects, the mesh size has to be determined consistently with, say, the renormalized electron (kinetic) energy calculated from the many-body theory [14]. This is of course not an easy task, especially for the case of small device structures. On the other hand, since the size of localized electrons in the MC device simulations is roughly given by the size of the mesh, this is not consistent with the concept of the electron wave packet. The BTE (or equivalently, the microscopic simulation) assumes that the electrons are localized and described by the wave packet whose size is comparable to the de Broglie wavelength. However, the size of the active device region is now comparable with the size of the wave packet in nanoscale MOSFETs and so it is not clear how the localized electrons in the channel should be interpreted in such microscopic simulations.

2 Inclusion of Quantum-Mechanical Size Quantization and Tunneling Effects in Particle-Based Device Simulators

2.1 Quantum-Mechanical Size Quantization Effects in Conjunction with Device Simulators

Successful scaling of MOSFETs towards shorter channel lengths requires thinner gate oxides and higher doping levels to achieve high drive currents and minimized short-channel effects [15, 16]. For these nanometer devices it was demonstrated a long time ago that, as the oxide thickness is scaled to 10 nm and below, the total gate capacitance is smaller than the oxide capacitance due to the comparable values of the oxide and the inversion layer capacitances. As a consequence, the device transconductance is degraded relative to the expectations of the scaling theory [17]. The inversion layer capacitance was also identified as being the main cause

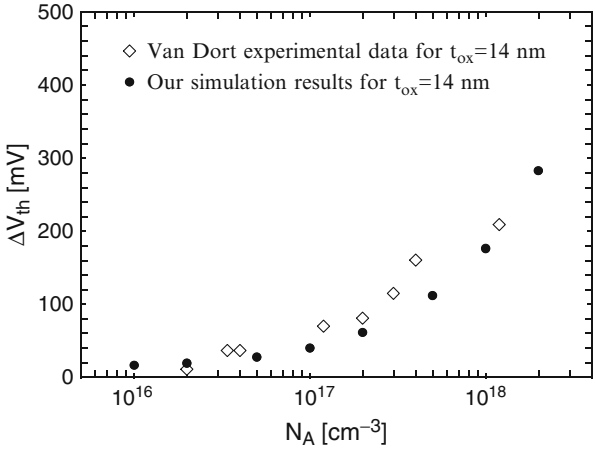


Fig. 2.6 SCHRED simulation data for the shift in the threshold voltage compared to the experimental values provided by van Dort and co-workers [20, 21]

of the second-order thickness dependence of MOSFET's *IV*-characteristics [18]. The finite inversion layer thickness was estimated experimentally by Hartstein and Albert [19]. The high levels of substrate doping, needed in nano-devices to prevent the punch-through effect has lead to quasi-two-dimensional (Q2D) nature of the carrier transport which is found responsible for the increased threshold voltage and decreased channel mobility, and a simple analytical model that accounts for this effect was proposed by van Dort and co-workers [20, 21]. Later on, Vasileska and Ferry [22] confirmed these findings by investigating the doping dependence of the threshold voltage in MOS capacitors. The experimental data for the doping dependence of the threshold voltage shift and our simulation results from [22] are shown in Fig. 2.6.

These results clearly demonstrate the influence of quantum-effects on the operation of nano-scale MOSFETs in both the off- and the on-state. The two physical origins of the inversion layer capacitance due to the finite density of states and due to the finite inversion layer thickness were demonstrated experimentally by Takagi and Toriumi [23]. A computationally efficient three-subband model that predicts both the quantum-mechanical effects in the electron inversion layer and the electron distribution within the inversion layer was proposed and implemented into the PICSEC simulator [24]. The influence of the image and many-body exchange-correlation effects on the inversion layer and the total gate capacitance was studied by Vasileska et al. [25]. It was also pointed out that the depletion of the poly-silicon gates considerably affects the magnitude of the total gate capacitance [26].

The above examples outline the advances during the two decades of research on the influence of quantum-effects on the operation on nano-devices. The conclusion is that any state-of-the-art device simulator must take into consideration the quantum-mechanical nature of the carrier transport and the poly-depletion effects to correctly predict the device off- and on-state behavior. As noted by many of

these authors, to account for the quantum-mechanical effects, one in principle has to solve the 2D/3D Schrödinger-Poisson problem in conjunction with an appropriate transport kernel. (For devices in which velocity overshoot is strongly pronounced, minimum that one can do is to solve the Boltzmann transport equation using the Ensemble Monte Carlo (EMC) technique.) Since the exact solution of the 2D/3D Schrödinger–Poisson problem is time-consuming even with present state-of-the-art computers, alternative paths have been sought for device simulators that utilize quantum potentials.

The idea of quantum potentials originates from the hydrodynamic formulation of quantum mechanics, first introduced by de Broglie and Madelung [27–29], and later developed by Bohm [30, 31]. In this picture, the wave function is written in complex form in terms of its amplitude $R(\mathbf{r},t)$ and phase $\psi(\mathbf{r},t) = R(\mathbf{r},t)\exp[iS(\mathbf{r},t)/\hbar]$. These are then substituted back into the Schrödinger equation to obtain the following coupled equations of motion for the density and phase

$$\frac{\partial \rho(\mathbf{r},t)}{\partial t} + \nabla \cdot \left(\rho(\mathbf{r},t) \frac{1}{m} \nabla S(\mathbf{r},t) \right) = 0, \tag{2.7}$$

$$-\frac{\partial S(\mathbf{r},t)}{\partial t} = \tfrac{1}{2m} \left[\nabla S(\mathbf{r},t) \right]^2 + V(\mathbf{r},t) + Q(\rho,\mathbf{r},t), \tag{2.8}$$

where $\rho(\mathbf{r},t) = R^2(\mathbf{r},t)$ is the probability density. By identifying the velocity as $\frac{1}{m}\nabla S$, and the flux as $\mathbf{j} = \rho \mathbf{v}$, (2.7) becomes the continuity equation. Hence, (2.7) and (2.8) arising from this so-called *Madelung transformation* to the Schrödinger equation have the form of classical hydrodynamic equations with the addition of an extra potential, often referred to as the *quantum* or *Bohm potential*, written as

$$V_Q = -\frac{\hbar^2}{2mR} \nabla^2 R \rightarrow -\frac{\hbar^2}{2m\sqrt{n}} \nabla^2 \sqrt{n} \tag{2.9}$$

where the density n is related to the probability density as $n(\mathbf{r},t) = N\rho(\mathbf{r},t) = NR^2(\mathbf{r},t)$, where N is the total number of particles. The Bohm potential essentially represents a field through which the particle interacts with itself. It has been used, for example, in the study of wave packet tunneling through barriers [32], where the effect of the quantum potential is shown to lower or smoothen barriers, and hence *allow* for the particles to leak through.

An alternate form of the quantum potential was proposed by Iafrate, Grubin and Ferry [33], who derived a form of the quantum potential based on moments of the *Wigner–Boltzmann equation*, the kinetic equation describing the time evolution of the Wigner distribution function [34]. Their form is based on moments of the Wigner function in the pure state, and involve an expansion of order $O(\hbar^2)$, which is given by

$$V_Q = -\frac{\hbar^2}{8m} \nabla^2 (\ln n), \tag{2.10}$$

this is sometimes referred to as the Wigner potential, or as the density gradient correction. Such quantum potentials have been extensively used in *density-gradient* and *quantum-hydrodynamic* methods. Their use in particle-based simulation schemes becomes questionable due to the presence of statistical noise in the representation of the electron density and the considerable difficulty to calculate the second derivative of the density on a completely unstructured mesh given by the particle discretization.

To avoid this problem, Ferry and Zhou derived a form for a smooth quantum potential [35], based on the effective classical partition function of Feynman and Kleinert [36]. More recently, Gardner and Ringhofer [37] derived a smooth quantum potential for hydrodynamic modeling, valid to all orders of $\hbar^2$, which involves a smoothing integration of the classical potential over space and temperature. There, it was shown that close to the equilibrium regime, the influence of the potential on the ensemble can be replaced by the classical influence of a smoothed non-local barrier potential. While this effective potential depends non-locally on the density, it does not directly depend on its derivatives. Through this effective quantum potential, the influence of the barriers on an electron is felt at quite some distance from the barrier. The smoothed effective quantum potential has been used successfully in quantum-hydrodynamic simulations of resonant tunneling effects in one-dimensional double-barrier structures [38].

In analogy to the smoothed potential representations discussed above for the quantum hydrodynamic models, it is desirable to define a smooth quantum potential for use in quantum particle-based simulations. Ferry [40] has suggested an *effective potential scheme* that emerges from a wave packet description of the particle motion, where the extent of the wave packet spread is obtained from the range of wavevectors in the thermal distribution function (characterized by an electron temperature). The effective potential, V_{eff}, is related to the self-consistent Hartree potential V, obtained from the Poisson equation, through an integral smoothing relation

$$V_{\text{eff}}(\mathbf{x}) = \int V(\mathbf{x}+\mathbf{y})\, G(\mathbf{y}, a_0)\, d\mathbf{y} \tag{2.11}$$

where G is a Gaussian with standard deviation a_0. The effective potential V_{eff} is then used to calculate the electric field that accelerates the carriers in the transport kernel of the Monte Carlo particle-based device simulator discussed in [39]. The calculation of V_{eff} has a fairly low computational cost, but the requirement that the electric field is updated every 0.01 fs to get physically accurate particle trajectories and to eliminate the artificial heating of the carriers in the vicinity of the Si/SiO_2 interface (where the fields are the strongest), adds to the computational cost. Note also that within this approach the parameter a_0 has to be adjusted in the initial stages of the simulation via comparisons of the sheet/line density of the Q2D/Q1D structure being investigated using the effective potential approach and the 1D/2D Schrödinger–Poisson simulations.

In this book chapter, in addition to the effective potential approach due to Ferry [40], we present a new form of the effective quantum potential for use in Monte Carlo device simulators. The proposed approach is based on perturbation

theory around thermodynamic equilibrium and leads to an effective potential which depends on the energy and wavevector of each individual electron, thus effectively lowering step-function barriers for high-energy carriers [41]. The quantum potential is derived from the idea that the Wigner and the Boltzmann equation with the quantum corrected potential should possess the same steady state. The resultant quantum potential is in general two-degrees smoother than the original Coulomb and barrier potentials, i.e. possesses two more classical derivatives which essentially eliminate the problem of statistical noise. The computation of the quantum potential involves only the evaluation of pseudo-differential operators and can therefore, be effectively facilitated using Fast Fourier Transform (FFT) algorithms. The approach is quite general and can easily be modified to modeling of, for example, triangular quantum wells. The above-described approach has been used in simulation of 25 nm MOSFET device with oxide thickness of 1.2 nm.

2.1.1 Thermodynamic Effective Potential

The basic idea of the thermodynamic approach to effective quantum potentials is that the resulting semiclassical transport picture should yield the correct thermalized equilibrium quantum state. Using quantum potentials, one generally replaces the quantum Liouville equation

$$\partial_t \rho + \frac{i}{\hbar}[H,\rho] = 0 \tag{2.12}$$

for the density matrix $\rho(x,y)$ by the classical Liouville equation

$$\partial_t f + \frac{\hbar}{2m^*} k \cdot \nabla_x f - \frac{1}{\hbar} \nabla_x V \cdot \nabla_k f = 0, \tag{2.13}$$

for the classical density function $f(x,k)$. Here, the relation between the density matrix and the density function f is given by the Weyl quantization,

$$f(x,k) = W[\rho] = \int \rho(x+y/2, x-y/2) \exp(ik \cdot y) dy. \tag{2.14}$$

The thermal equilibrium density matrix in the quantum mechanical setting is given by $\rho^{eq} = \mathrm{e}^{-\beta H}$, where $\beta = 1/k_B T$ is the inverse energy and the exponential is understood as a matrix exponential, i.e. $\rho^{eq}(x,y) = \sum_\lambda \psi_\lambda(x) \exp(-\beta\lambda) \psi_\lambda(y)^*$ holds with $\{\psi_\lambda\}$ the orthonormal eigensystem of the Hamiltonian H. On the other hand, in the semiclassical transport picture, the thermodynamic equilibrium density function f_{eq} is given by the Maxwellian $f_{eq}(x,k) = \exp\left(-\frac{\beta\hbar^2|k|^2}{2m^*} - \beta V\right)$. Consequently, to obtain the quantum mechanically correct equilibrium states in the semiclassical Liouville equation with the effective quantum potential V^Q, we set

$$\begin{aligned} f_{eq}(x,k) = \exp\left(-\frac{\beta\hbar^2|k|^2}{2m^*} - \beta V^Q\right) &= W[\rho^{eq}] \\ &= \int e^{-\beta H} \rho(x+y/2, x-y/2) \exp(ik \cdot y) dy. \end{aligned} \tag{2.15}$$

This basic concept was originally introduced by Feynman and Kleinert [36]. Different forms of the effective quantum potential arise from different approaches to approximate the matrix exponential $e^{-\beta H}$.

In the approach presented in this paper, we represent $e^{\beta H}$ as the Green's function of the semigroup generated by the exponential. Introducing an artificial dimensionless parameter γ and defining $\rho(x,y,\gamma) = \sum_\lambda \psi_\lambda(x)\exp(-\gamma\beta\lambda)\psi_\lambda(y)^*$, we obtain a heat equation for ρ by differentiating ρ w.r.t. γ and using the eigenfunction property of the wave functions ψ_λ. This heat equation is referred to as the Bloch equation

$$\partial_\gamma \rho = -\frac{\beta}{2}\left(H\cdot\rho + \rho\cdot H\right),\ \rho(x,y,\gamma=0) = \delta(x-y), \tag{2.16}$$

and $\rho^{eq}(x,y)$ is given by $\rho(x,y,\gamma=1)$. Under the Weyl quantization this becomes with the usual Hamiltonian $H = -\frac{\hbar^2}{2m^*}\Delta_x + V$ and defining the effective energy E by $f = W[\rho] = e^{-\beta E}$,

$$\begin{aligned}\partial_\gamma E = {} & \frac{\beta\hbar^2}{8m^*}\left(\Delta_x E - \beta|\nabla_x E|^2\right) + \frac{\hbar^2|k|^2}{2m^*} \\ & + \frac{1}{2(2\pi)^3}\sum_{\nu=\pm1}\int V(x+\nu y/2)\exp[\beta E(x,k,\gamma) - \beta E(x,q,\gamma) \\ & + iy(k-q)]dqdy, E(x,k,\gamma=0) = 0.\end{aligned} \tag{2.17}$$

The effective quantum potential in this formulation is given by $E(x,k,\gamma=1) = V^Q + \frac{\hbar^2|k|^2}{2m^*}$. The logarithmic Bloch equation is now solved 'asymptotically' using the *Born approximation*, i.e. by iteratively inverting the highest order differential operator (the Laplacian). This involves successive solution of a heat equation for which the Green's function is well known, giving (see [42] for the details),

$$V^Q(x,k) = \frac{1}{(2\pi)^3}\int \frac{2m^*}{\beta\hbar^2 k\cdot\xi}\sinh\left(\frac{\beta\hbar^2 k\cdot\xi}{2m^*}\right)\exp\left(-\frac{\beta\hbar^2}{8m^*}|\xi|^2\right)V(y)e^{i\xi\cdot(x-y)}dyd\xi. \tag{2.18}$$

Note that the effective quantum potential V^Q now depends on the wave vector k. For electrons at rest, i.e. for $k=0$, the effective potential V^Q reduces to the Gaussian smoothing given in (2.11) and [40]. Also note that there are no fitting parameters in this approach, i.e. the size of the wavepacket is determined by the particle's energy.

The potential $V(y)$ that appears in the integral of (2.18) can be represented as a sum of two potentials: the barrier potential $V_B(x)$, which takes into account the discontinuity at the Si/SiO_2 interface due to the difference in the semiconductor and the oxide affinities and the Hartree potential $V_H(x)$ that results from the solution of the Poisson equation. Note that the barrier potential is 1D and independent of time and needs to be computed only once in the initialization stage of the code. On the other hand, the Hartree potential is 2D and time-dependent it describes the evolution of charge from quasi-equilibrium to a non-equilibrium state. Since the evaluation

of the effective Hartree potential as given by (2.18), is very time consuming and CPU intensive, approximate solution methods have been pursued to resolve this term within a certain level of error tolerance.

We recall from the above discussion that the barrier potential is just a step-function. Under these circumstances $e\nabla_x V_B(x) = B(1,0,0)^T \delta(x_1)$, where B is the barrier height (in the order of 3.2 eV) and x_1 is a vector perpendicular to the interface. We actually need only the gradient of the potential so that using the pseudo-differential operators, we compute

$$\nabla_x V_B^Q(x,p) = \exp\left[\frac{\beta\hbar^2|\nabla_x|^2}{8m^*}\right] \frac{2m^* \sin\left(\frac{\beta\hbar p\cdot\nabla_x}{2m^*}\right)}{\beta\hbar p\cdot\nabla_x} \nabla_x V_B(x). \tag{2.19}$$

This gives

$$e\nabla_x V_B^Q(x,p) = \frac{B}{2\pi}(1,0,0)^T \int \exp\left[-\beta\frac{\hbar^2|\xi_1|^2}{8m^*}\right] \frac{2m^* \sinh\left(\frac{\beta\hbar p_1\cdot\xi_1}{2m^*}\right)}{\beta\hbar p_1\cdot\xi_1} e^{i\xi_1\cdot x_1} d\xi_1 \tag{2.20}$$

Note that V_B^Q is only a function of (x_1,p_1), i.e. it remains to be strictly one-dimensional, where x_1 and p_1 are the position and the momentum vector perpendicular to the interface. This when combined with the fact that we have to calculate this integral only once is a reason why we have decided to tabulate the result given by (2.20) on a mesh.

The Hartree potential, as computed by solving the d-dimensional Poisson equation depends in general upon d particle coordinates. For example, on a rectangular mesh the 2D Hartree potential is given by $V_H(x_1,x_2,t)$, and one has to evaluate $V_H^Q(x_1,x_2,p_1,p_2,t)$ using (2.18) N times each time step for all particles position and momenta: $x^n, p^n, n = 1,\ldots,N$ (where N is the number of electrons, which is large). Of course, this is an impossible task to be accomplished in finite time on present state-of-the-art computers. We, therefore, suggest the following scheme. According to (2.18), we evaluate the quantum potential by multiplying the Hartree potential by a function of $\hbar\nabla_x$, or by multiplying the Fourier transform of the Hartree potential by a function of $\hbar\xi$. We factor the expression in (2.18) into

$$\begin{aligned} V_H^Q(x,k) &= \frac{2im^*}{\beta\hbar^2 k\cdot\nabla_x} \sinh\left(\frac{\beta\hbar^2 k\cdot\nabla_x}{2im^*}\right) \exp\left(\frac{\beta\hbar^2}{8m^*}|\nabla_x|^2\right) V_H(x) \\ &= \frac{2im^*}{\beta\hbar^2 k\cdot\nabla_x} \sinh\left(\frac{\beta\hbar^2 k\cdot\nabla_x}{2im^*}\right) V_H^0(x), \end{aligned} \tag{2.21}$$

with

$$V_H^0(x) = \exp\left(\frac{\beta\hbar^2}{8m^*}|\nabla_x|^2\right) V_H(x). \tag{2.22}$$

The evaluation of the potential $V_H^0(x)$, which is a version of the Gaussian smoothed potential due to Ferry [40]. This is computationally inexpensive since it does not depend on the wavevector k. On the other hand because of the Gaussian smoothing, $V_H^0(x)$ will be a smooth function of position, even if the Hartree potential $V_H(x)$ is computed via the Poisson equation where the electron density is given by a particle discretization. Therefore, the Fourier transform of the potential $V_H^0(x)$ will decay rapidly as a function of ξ, and it is admissible to use a Taylor expansion for small values of $\hbar\xi$ in the rest of the operator. This gives

$$\frac{2im^*}{\beta\hbar^2 k\cdot\nabla_x}\sinh\left(\frac{\beta\hbar^2 k\cdot\nabla_x}{2im^*}\right)\approx 1-\frac{\beta^2\hbar^4(k\cdot\nabla_x)^2}{24(m^*)^2}, \tag{2.23}$$

or

$$\partial_{x_r}V_H^Q(x^n,p^n)=\partial_{x_r}V_H^0(x^n)-\frac{\beta^2\hbar^2}{24m^{*2}}\sum_{j,k=1}^{2}p_j^n p_k^n\partial x_j\partial x_k\partial x_r V_H^0(x^n),\quad n=1,\ldots,N \tag{2.24}$$

for all particles. This is done simply by numerical differentiation of the sufficiently smooth grid function V_H^0 and interpolation. The evaluation of (2.24) is the price we have to pay when we compare the computational cost of this approach as opposed to the Ferry approach [40] which uses simple forward, backward or centered difference scheme for the calculation of the electric field. However, with this novel effective potential approach we avoid the use of adjustable parameters.

Example: Quantum Effects in a Conventional 25 nm MOSFET

As a first example to which we apply the Ringhofer's effective potential approach we take conventional MOSFET device with 25 nm channel length. The parameters of the device structure being simulated are as follows: the average channel/substrate doping is $10^{19}\,\mathrm{cm}^{-3}$, the doping of the source and drain regions is $10^{19}\,\mathrm{cm}^{-3}$, the junction depth is 30 nm, the oxide thickness is 1.2 nm and the gates are assumed to be metal gates with work-function equal to the semiconductor affinity. The gate/channel length is 25 nm. First in Fig. 2.7, the carrier confinement within the triangular potential well with and without the inclusion of the quantum-mechanical size-quantization effects is shown for the bias conditions $V_G = V_D = 1$V. From the results shown in this figure, it is evident that the low-energy electrons are displaced little more than the high-energy electrons; the reason being the fact that the high-energy electrons tend to behave as classical particles and hence are displaced relatively less. Also note that there is practically no carrier heating for the case when the effective potential is used in calculating the driving electric field. The carrier displacement from the interface proper is also seen from the results presented in Fig. 2.8. Notice that there is approximately 2 nm average shift of the electron density distribution near the source end of the channel when quantization effects are included in the model.

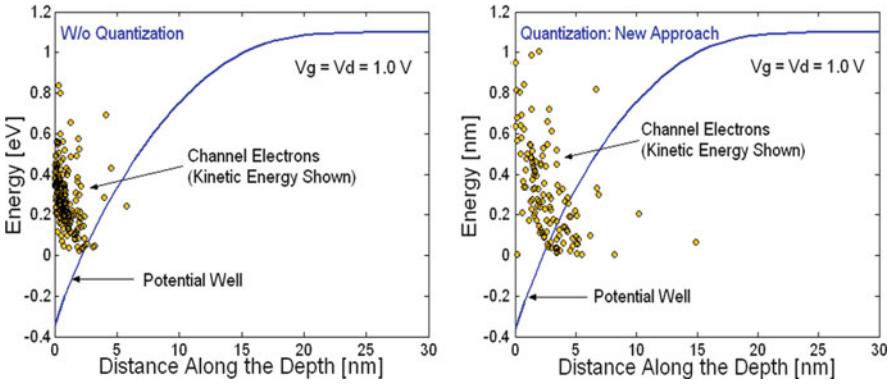


Fig. 2.7 Electron localization within the triangular potential barrier for the case when quantization effects are not included in the model *(left panel)* and for the case when we include quantum-mechanical space-quantization effects by using the effective potential approach presented in this paper *(right panel)*. The potential profile is taken in the middle portion of the channel, not at the drain end, and because of that some electrons seem to be in regions where they should not, but that is just an artifact of presenting the results. The triangular potential at the drain end of the channel is much wider

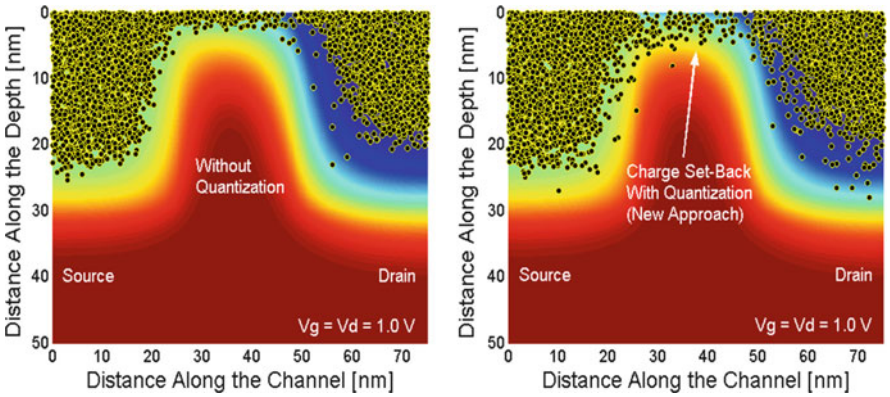


Fig. 2.8 Electron distribution in the device without *(left panel)* and with *(right panel)* the incorporation of quantum-mechanical size-quantization effects

Also note that carriers behave more like bulk carriers at the drain end of the channel and are displaced in the same manner when using both the classical and the quantum-mechanical model.

The channel length variation of the sheet electron density is shown in Fig. 2.9 for classical, fully-quantum ($V_H^Q + V_B^Q$) and quantum-barrier field (V_B^Q) models [43]. Also compared are the simulation results for the sheet electron density from the new method with those utilizing the approach due to Ferry [44]. There are several noteworthy features to be observed in this figure. First, the pinch-off of the sheet

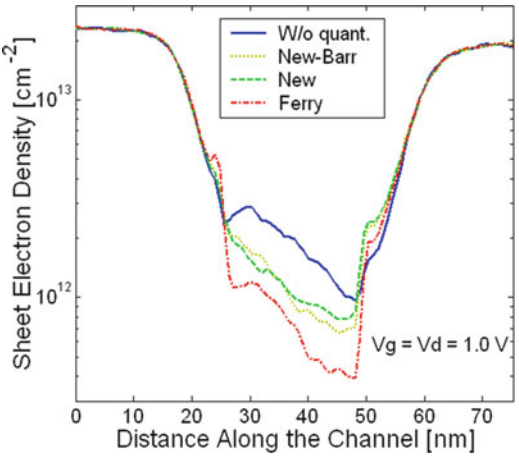


Fig. 2.9 Variation of the sheet electron density along the channel. *New-barr* corresponds to the case when we only include the influence of the barrier field. *New* represents the case when we include both the barrier and the Hartree contributions to the total electric field

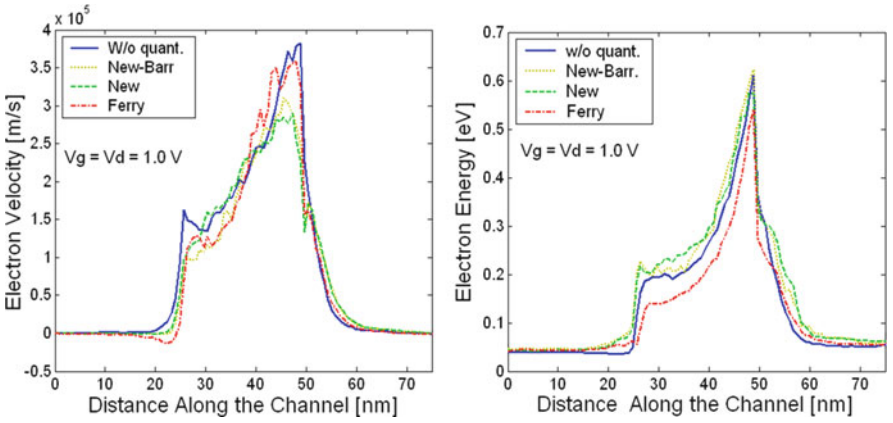


Fig. 2.10 Average electron velocity *(left panel)* and average electron energy *(right panel)* variation along the channel

electron density near the drain end of the channel is evident in all models used. Second, the barrier and the full-effective potential scheme give almost the same value for the sheet electron density, which suggests that the repulsive barrier field dominates over the attractive field due to the Hartree potential. Third, the method due to Ferry leads to significantly lower value for the sheet electron density which can be improved by choosing lower values of the Gaussian smoothing parameter.

The average electron velocity and the average electron energy are shown in the left and the right panels of Fig. 2.10, respectively. Comparing the results for the average carrier energy on the right panel, one can see that the data for the case when one has not included the effective potential and the case when one has used the new model for the effective potential agree very well with each other. The slight increase in the carrier energy in the channel region (which is non-physical) when

one uses the new effective potential approach is because of the very high value of the quantum field being present in the vicinity of the Si/SiO_2 interface proper. The situation can be improved by using a sufficiently small time-step (for example 0.01 fs) during Monte Carlo simulation. The approach due to Ferry gives significantly lower value for the carrier energy near the source end of the channel which has been explained to be due to the bandgap widening effect. Also, here we do not observe the non-physical carrier heating because of the fact that Ferry's effective potential is calculated from the mesh potential which depends on both the meshing and the Gaussian parameter used in the model. The quantum field is calculated from direct differentiation of the effective mesh-potential and has every possibility of being underestimated due to the finite size of the meshing used in simulations. It also is independent on carrier energy (according to the current implementation of the model). When one confronts these data with the results for the average electron velocity, its east to say that in the low-energy region near the source end of the channel the velocity is almost the same for all cases considered. At the drain end, one finds degradation of the velocity due to the smearing introduced by the quantum potential. Again, the inclusion of the barrier field and of the quantum-corrected Hartree term give similar values, which suggests that for the device being considered in this study only the barrier field has significant impact [45].

The device transfer characteristics are shown in the left panel of Fig. 2.11. Again, it becomes clear that the proposed full quantum potential and the barrier potential give similar values for the current. Looking more in detail the device transfer characteristics one finds that the quantization effects lead to threshold voltage increase of about 220 mV. When properly adjusted for the oxide thickness difference, this result is consistent with previously published data [20]. Evidently, as deduced from the output characteristics shown in the right panel of Fig. 2.11, the shift in the threshold voltage leads to a decrease in the on-state current by 30%. The later observation confirms earlier findings that one must include quantum effects into the theoretical model to be able to properly predict the device threshold voltage and its on-state current.

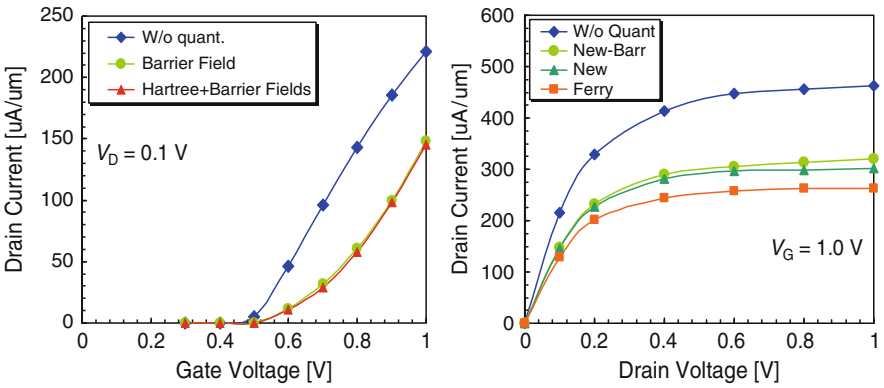


Fig. 2.11 Device transfer characteristic for $V_D = 0.1$ V *(left panel)*. Device output characteristics for $V_G = 1.0$ V *(right panel)*

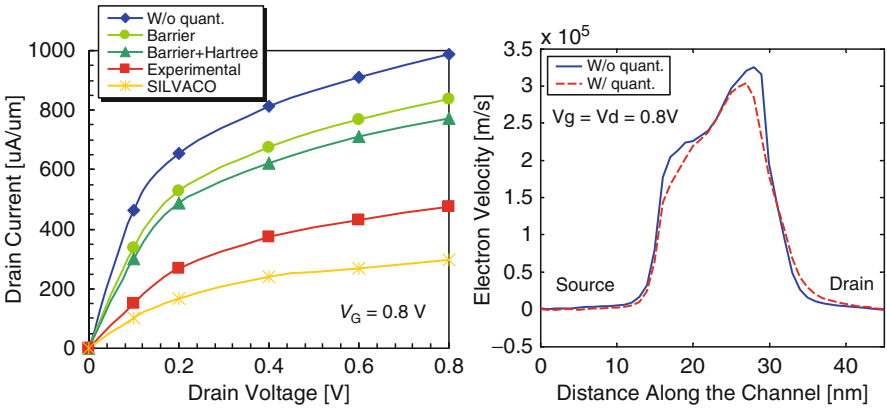


Fig. 2.12 *Left panel*: Conventional 15 nm MOSFET device output characteristics. *Right panel*: Average electron velocity along the channel

Next, the simulation results of a 15 nm conventional *n*-channel MOSFET device are discussed. Similar devices have been fabricated by Intel Corporation [46]. The physical gate length of the device used is 15 nm. The source/drain length equals 15 nm and the junction depth is also 15 nm. The bulk substrate thickness used for simulations is 45 nm. The height of the fabricated polysilicon gate electrode for this device is 25 nm. The gate oxide used was SiO_2 with physical thickness of only 0.8 nm. The source/drain doping density is $2 \times 10^{19}\,\mathrm{cm}^{-3}$ and the channel doping is $1.5 \times 10^{19}\,\mathrm{cm}^{-3}$. The substrate doping used is $1 \times 10^{18}\,\mathrm{cm}^{-3}$. The simulated device output characteristics are shown in Fig. 2.12.

There are again several noteworthy features in these results: (1) Quantum-mechanical size quantization increases the threshold voltage as observed from the decrease in the slope in the linear region and hence degrades the device transconductance. (2) Drain current degradation due to the quantum effects is not uniform rather decreases with the increase in drain bias. The reason may be attributed again to the fact that the electrons tend to behave as classical particles as average carrier energy increases with the increase in drain bias, (3) there is a considerable difference between the barrier-correction and the barrier-Hartree (full) correction which is mainly due to the use of higher doping density $(1.5 \times 10^{19}\,\mathrm{cm}^{-3})$ in the channel region than was used in the 25 nm MOSFET $(1 \times 10^{19}\,\mathrm{cm}^{-3})$ case. The higher doping density has a direct impact on the Hartree potential making the triangular channel potential steeper and hence introducing a pronounced quantum effects. But the overall degradation of the drain current as compared to the 25 nm MOSFET device structure has reduced in the 15 nm device because of the ballistic nature of the carrier motion in the latter case. This fact becomes clear if one observes the velocity profile of the device as depicted in the right panel of Fig. 2.12. What is important in this figure is that the carriers attain a velocity which is comparable to that in the 25 nm device structure even with a lesser biases applied i.e. $V_G = V_D = 0.8\,\mathrm{V}$. Also, the gate oxide thickness is lesser in the 10 nm device which means that the gate oxide capacitance

constitutes the major portion of the total effective gate capacitance thereby reducing the impact of the quantum capacitance. (4) The discrepancy between the experimental and the simulated results is attributed mainly to two reasons: (a) the series resistance coming from the finite width of the actual device structure and the contact resistances, and (b) the gate polysilicon depletion effects which as previously mentioned, can introduce further degradation of the drain current on the order of 10–30% depending on the doping density and the height of the polysilicon gate used. The limited data as supplied by the Intel Corporation shows that the polysilicon gate is of 25 nm height which can indeed contribute to a significant degradation of the drain current. (5) The use of a commercial simulator like the drift-diffusion based SILVACO Atlas fails considerably to predict the device behavior mainly because of the ballistic and quantized nature of the carriers in these nanoscale device structures.

Example: Size-Quantization in Nanoscale SOI Devices

Because of using lightly/nearly undoped channel region, size-quantization effects in nanoscale fully-depleted SOI devices find a major source in the very physical nature of the confined region which remains sandwiched between the two oxide layers. In order to verify the applicability of the quantum potential approach developed in this work, a single gated SOI device structure will be studied first. Simulations will be carried out to calculate the threshold voltage as a function of the silicon film thickness and the results will be compared to other available methods. The SOI device used here has the following specifications: gate length is 40 nm, the source/drain length is 50 nm each, the gate oxide thickness is 7 nm with a 2 nm source/drain overlap, the box oxide thickness is 200 nm, the channel doping is uniform at $1 \times 10^{17}\,\mathrm{cm}^{-3}$, the doping of the source/drain regions equals $2 \times 10^{19}\,\mathrm{cm}^{-3}$, and the gate is assumed to be a metal gate with workfunction equal to the semiconductor affinity. There is a 10 nm spacer region between the gate and the source/drain contacts. The silicon (SOI) film thickness is varied over a range of 1–10 nm for the different simulations that were performed to capture the trend in the variations of the device threshold voltage. Similar experiments were performed in [47, 48] using the Schrödinger–Poisson solver and Ferry's effective potential approaches, respectively. For comparison purposes, threshold voltage is extracted from the channel inversion density vs. gate bias profile and extrapolating the linear region of the characteristics to a zero value. This method also corresponds well to the linear extrapolation technique using the drain current-gate voltage characteristics.

The results showing the trend in the threshold voltage variation with respect to the SOI film thickness are depicted in Fig. 2.13. One can see that Ferry's effective potential approach overestimates the threshold voltage for a SOI thickness of 3 nm due to the use of a rather approximate value for the standard deviation of the Gaussian wave packet which results in a reduced sheet electron density. As the silicon film thickness decreases, the resulting confining potential becomes more like rectangular from a combined effects of both the inversion layer quantization and the SOI film (physical) quantization, which also emphasizes the need for using a more

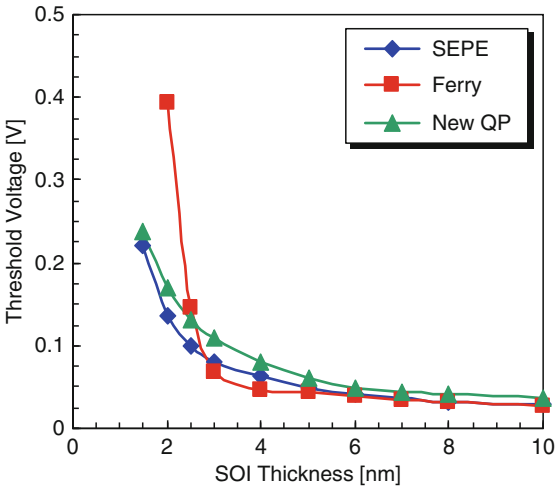


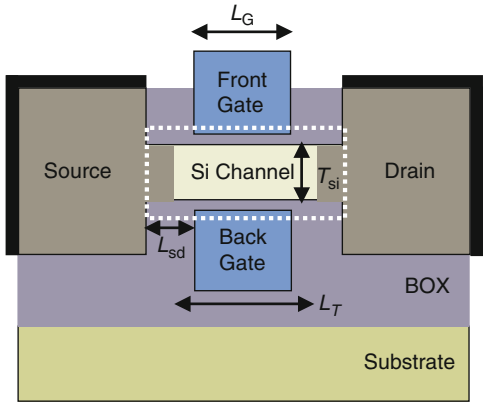
Fig. 2.13 Threshold voltage variation with SOI film thickness. SEPE stands for Schrödinger-Poisson, Ferry stands for Ferry's effective potential approach and New QP stands for new quantum potential

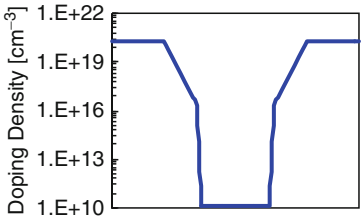
realistic quantum-mechanical wavepacket description for the confined electrons. Of most importance in this figure is the very fact that the new quantum potential approach is free from this large discrepancy and can capture the trend in the threshold voltage as it is obtained from the more accurate 2D Schrödinger-3D Poisson solver. These results indicate that the new quantum potential method can be applied to the simulations of SOI devices with a greater accuracy and predictive capability as it will be seen from the results presented in the next section.

Example: Size-Quantization in Nanoscale DG SOI Devices

Figure 2.14 shows the simulated DG SOI device structure used in this work, which is similar to the devices reported in [49]. For quantum simulation purposes only the dotted portion of the device, termed as the *intrinsic* device is taken into considerations. The device was originally designed in order to achieve the ITRS performance specifications for the year 2016.

The effective intrinsic device consists of two gate stacks (gate contact and SiO_2 gate dielectric) above and below a thin silicon film. For the intrinsic device, the thickness of the silicon film is 3 nm. Use of a thicker body reduces the series resistance and the effect of process variation but it also degrades the short channel effects (SCE). From the SCE point of view, a thinner body is preferable but it is harder to fabricate very thin films of uniform thickness, and the same amount of process variation ($\pm 10\%$) may give intolerable fluctuations in the device characteristics. A thickness of 3 nm seems to be a reasonable compromise, but other body thicknesses are also examined. The top and bottom gate insulator thickness is 1 nm, which is expected to be near the scaling limit for SiO_2. As for the gate contact, a metal gate with tunable workfunction, Φ_G, is assumed, where Φ_G is adjusted to





T_{ox} = 1 nm
L_G = 9 nm
L_{sd} = 10 nm
N_b = 0
Φ_G = 4.188

T_{si} = 3 nm
L_T = 17 nm
N_{sd} = 2 x 10^{20} cm^{-3}
g = 1 nm/decade
V_G = 0.4 V

Fig. 2.14 DG device structure being simulated

4.188 eV to provide a specified off-current value of $4\,\mu A/\mu m$. The background doping of the silicon film is taken to be intrinsic, however due to diffusion of the dopant ions, the doping profile from the heavily doped S/D extensions to the intrinsic channel is graded with a coefficient of g which equals to 1 nm/dec. For convenience, the doping scheme is also shown in Fig. 2.14. According to the roadmap, the high performance (HP) device should have a gate length of $L_G = 9\,nm$ at the year 2016. At this scale, two-dimensional (2D) electrostatics and quantum mechanical effects both play an important role and traditional device simulators may not provide reliable projections. The length L_T, is an important design parameter in determining the on-current, while gate metal workfunction Φ_G, directly controls the off-current. The doping gradient g, affects both on-current and off-current. Values of all the structural parameters of the device are shown in Fig. 2.14 as well.

The intrinsic device is simulated using the new quantum potential approach in order to gauge the impact of size-quantization effects on the DG SOI performance. The results are then compared to that from a full quantum approach based on the non-equilibrium Green's function (NEGF) formalism (NanoMOS–2.5) developed

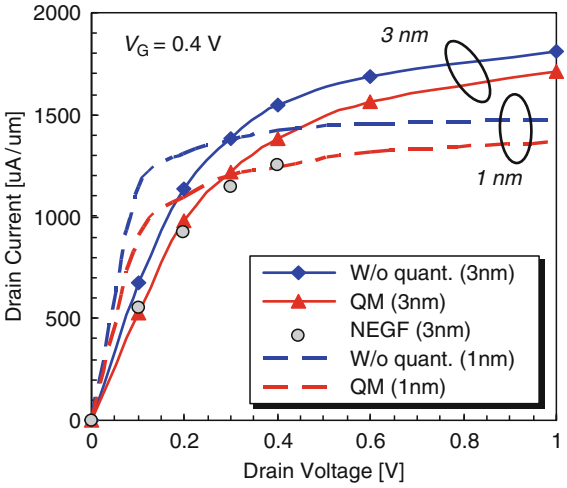


Fig. 2.15 Generic DG SOI device output characteristics

at Purdue University [50]. In this method, scattering inside the intrinsic device is treated by a simple Büttiker probe model, which gives a phenomenological description of scattering and is easy to implement under the Greens' function formalism. The simulated output characteristics are shown in Fig. 2.15. Devices with both 3 and 1 nm channel thickness are used with applied gate bias of 0.4 V. The salient features of this figure are as follows: (1) Even with an undoped channel region, the devices achieve a significant improvement with respect to the SCEs as depicted in flatness of the saturation region. This is due to the use of the two gate electrodes and an ultrathin SOI film which makes the gates gain more control on the channel charge. (2) Reducing the channel SOI film thickness to 1 nm further reduces the SCEs and improves the device performance. However, the reduction in the drive current at higher drain biases is due to series resistance effect pronounced naturally when the drain current increases. (3) Regarding the quantum effects, one can see that quantum-mechanical size quantization does not play a very dominant role in degrading the device drive current mainly because of use of an undoped channel region. Also, looking at the 3 nm (or 1 nm) case alone one can see that the impact of quantization effects reduces as the drain voltage increases because of the growing bulk nature of the channel electrons. (4) Percentage reduction in the drain current is more pronounced in 1 nm case throughout the range of applied drain bias because of the stronger physical confinement arising from the two SiO_2 layers sandwiching the silicon film. (5) Finally, the comparison between the quantum potential formalism and the NEGF approach for the device with 3 nm SOI film thickness shows reasonable agreement which further establishes the applicability of this method in the simulations of different technologically viable nanoscale classical and non-classical MOSFET device structures.

2.2 SCHRED First and Second Generation

Proper inclusion of the quantum-mechanical size quantization effects in device simulators is achieved by solving the Schrödinger–Poisson–Boltzmann problem. This approach was discussed in details in [51]. Here we only focus on solving the 1D Schrödinger–Poisson problem for proper description of charge quantization in MOS capacitors. This can be achieved with SCHRED First Generation tool that is installed on the Network for Computational Nanotechnology (www.nanoHUB. org). However, in the past 2–3 years many users of the existing SCHRED expressed wishes for increasing the present capabilities of SCHRED tool in terms of making it capable to study MOS capacitors made of silicon or strained silicon with arbitrary crystallographic transport directions and to be able to simulate MOS capacitors fabricated of other materials. To satisfy user needs, an effort was undertaken at ASU and SCHRED Second Generation was developed that has all the required features that were on the wish list of SCHRED First Generation. The tool was developed by a M.S. student of Prof. Vasileska at Arizona State University Gokula Kannan. In what follows, we will first explain the capabilities of the SCHRED First Generation Tool and then we will describe SCHRED Second Generation Tool in details.

2.2.1 SCHRED First Generation Capabilities

The periodic crystal potential in the bulk of semiconducting materials is such that, for a given energy in the conduction band, the allowed electron wavevectors trace out a surface in **k**-space. In the effective-mass approximation for silicon, these constant energy surfaces can be visualized as six equivalent ellipsoids of revolution (Fig. 2.16), whose major and minor axes are inversely proportional to the effective masses. A collection of such ellipsoids for different energies is referred to as a valley.

In this framework, the bulk Hamiltonian for an electron, residing in one of these valleys is of the form

$$H_o(\mathrm{R}) = -\left(\frac{\hbar^2}{2m_x^*}\frac{\partial^2}{\partial x^2} + \frac{\hbar^2}{2m_y^*}\frac{\partial^2}{\partial y^2} + \frac{\hbar^2}{2m_z^*}\frac{\partial^2}{\partial z^2}\right) + V_{eff}(z) = H_{o||}(\mathrm{r}) + H_{o\perp}(z), \tag{2.25}$$

where $\mathbf{R} = (\mathbf{r}, z)$, $V_{eff}(z) = V_H(z) + V_{exc}(z)$ is the effective potential energy profile of the confining potential, $V_H(z)$ is the Hartree potential which is nothing more but a solution of the 1D Poisson equation introduced later in the text, $V_{exc}(z)$ is the exchange-correlation potential also discussed later in the text, $H_{o||}$ is the parallel part of H_o, and the transverse part is defined as

$$H_{o\perp}(z) = -\frac{\hbar^2}{2m_z^*}\frac{\partial^2}{\partial z^2} + V_{eff}(z). \tag{2.26}$$

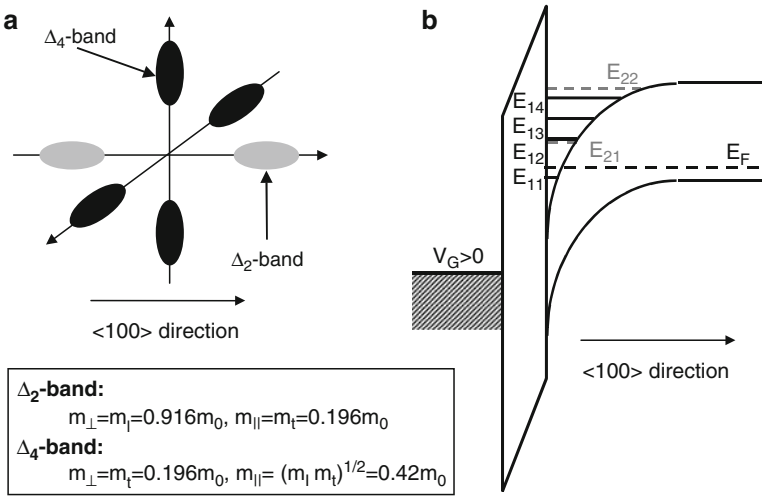


Fig. 2.16 *Right panel* – Potential diagram for inversion of *p*-type semiconductor. In this first notation E_{ij} refers to the jth subband from either the Δ_2-band (i = 1) or Δ_4-band (i = 2). *Left panel* – Constant-energy surfaces for the conduction-band of silicon showing six conduction-band valleys in the <100> direction of momentum space. The band minima, corresponding to the centers of the ellipsoids, are 85% of the way to the Brillouin-zone boundaries. The long axis of an ellipsoid corresponds to the longitudinal effective mass of the electrons in silicon, $m_l = 0.916m_o$, while the short axes correspond to the transverse effective mass, $m_t = 0.190m_o$. For <100> orientation of the surface, the Δ_2-band has the longitudinal mass (m_l) perpendicular to the semiconductor interface and the Δ_4-band has the transverse mass (mt) perpendicular to the interface. Since larger mass leads to smaller kinetic term in the Schrödinger equation, the unprimed ladder of subbands (as is usually called), corresponding to the Δ_2-band, has the lowest ground state energy. The degeneracy of the unprimed ladder of subbands for <100> orientation of the surface is 2. For the same reason, the ground state of the primed ladder of subbands corresponding to the Δ_4-band is higher than the lowest subband of the unprimed ladder of subbands, The degeneracy of the primed ladder of subbands for (100) orientation of the interface is 4

The basis-states of the unperturbed Hamiltonian are assumed to be of the form

$$\Psi_n(\mathrm{R}) = \frac{1}{\sqrt{A}} e^{i\mathrm{k}\cdot\mathrm{r}} \psi_n(z), \tag{2.27}$$

where **k** is a wavevector in the *xy*-plane and *A* is the area of the sample interface. The subband wavefunctions satisfy the one-dimensional Schrödinger equation,

$$H_{o\perp}(z)\psi_n(z) = \varepsilon_n \psi_n(z) \tag{2.28}$$

subject to the boundary conditions that $\psi_n(z)$ are zero for $z = 0$ and approach zero as $z \to \infty$. In (2.28), ε_n is the subband energy and $\psi_n(z)$ is the corresponding wavefunction. In the parabolic band approximation, the total energy of the electrons is given by

$$E_n(\mathrm{k}) = \frac{\hbar^2 \mathrm{k}^2}{2m^*_{xy}} + \varepsilon_n = \varepsilon_{\mathrm{k}} + \varepsilon_n, \tag{2.29}$$

where $\varepsilon_{\mathbf{k}}$ is the kinetic energy and m^*_{xy} is the density of states mass along the xy-plane. An accurate description of the charge in the inversion layer of deep-submicrometer devices and, therefore, the magnitude of the total gate capacitance C_{tot} requires a self-consistent solution of the 1D Poisson

$$\frac{\partial}{\partial z}\left[\varepsilon(z)\frac{\partial \phi}{\partial z}\right] = -e[N_D^+(z) - N_A^-(z) + p(z) - n(z)], \tag{2.30}$$

and the 1D Schrödinger equation

$$\left[-\frac{\hbar^2}{2m_i^\perp}\frac{\partial^2}{\partial z^2} + V_{eff}(z)\right]\psi_{ij}(z) = E_{ij}\psi_{ij}(z). \tag{2.31}$$

In (2.30) and (2.31), $\varphi(z)$ is the electrostatic potential [the Hartree potential $V_H(z) = -e\varphi(z)$], $\varepsilon(z)$ is the spatially dependent dielectric constant, $N_D^+(z)$ and $N_A^-(z)$ are the ionized donor and acceptor concentrations, $n(z)$ and $p(z)$ are the electron and hole densities, $V_{eff}(z)$ is the effective potential energy term that equals the sum of the Hartree and exchange-correlation corrections to the ground state energy of the system, $m_i^\perp$ is the effective mass normal to the semiconductor-oxide interface of the ith valley, and E_{ij} and $\psi_{ij}(z)$ are the energy level and the corresponding wavefunction of the electrons residing in the jth subband from the ith valley. The electron-density is calculated using

$$n(z) = \sum_{i,j} N_{ij}\psi_{ij}^2(z) \tag{2.32}$$

where N_{ij} is the sheet electron concentration in the ith subband from the jth valley is given by

$$N_{ij} = g_i\frac{m^*_{xy}}{\pi\hbar^2}k_BT\ln\left\{1+\exp[(E_F - E_{ij})/k_BT]\right\} \tag{2.33}$$

where g_i is the valley degeneracy factor and E_F is the Fermi energy. When evaluating the exchange-correlation corrections to the chemical potential, we have relied on the validity of the density functional theory (DFT) of Hohenberg and Kohn [52], and Kohn and Sham [53]. According to DFT, the effects of exchange and correlation can be included through a one-particle exchange-correlation term $V_{exc}[n(z)]$, defined as a functional derivative of the exchange-correlation part of the ground-state energy of the system with respect to the electron density $n(z)$. In the local density approximation (LDA), one replaces the functional $V_{exc}[n(z)]$ with a function $V_{exc}[n(z)] = \mu_{exc}[n_0 = n(z)]$, where μ_{exc} is the exchange-correlation contribution to the chemical potential of a homogeneous electron gas of density n_0, which is taken to be equal to the local electron density $n(z)$ of the inhomogeneous system. In our model, we use the LDA and approximate the exchange-correlation potential energy term $V_{exc}(z)$ by an interpolation formula developed by Hedin and Lundqvist [54]

$$V_{exc}(z) = -\frac{e^2}{8\pi\varepsilon_{sc}b}\left[1 + 0.7734x\ln\left(1+\frac{1}{x}\right)\right]\left(\frac{2}{\pi\alpha r_s}\right), \tag{2.34}$$

which is accurate over a large density range. In (2.34), $\alpha = (4/9\pi)^{1/3}$, $x = x(z) = r_s/21$, $r_s = r_s(z) = [4\pi b^3 n(z)/3]^{-1/3}$, and $b = 4\pi\varepsilon_{sc}\hbar^2/m^* e^2$. Exchange and correlation effects tend to lower the total energy of the system and lead to non-uniform shift of the energy levels and repopulation of the various subbands. The enhancement of the exchange-correlation contribution to the energy predominantly affects the ground subband of the occupied valley; the unoccupied subbands of the same valley are essentially unaffected. As a result, noticeable increase in the energy of the inter-subband transitions can be observed at high electron densities.

Similarly, the valence band is represented by the heavy hole band and light hole band, the spit-off band is ignored because the spit-off energy is large enough to exclude any hole staying there. In treating holes quantum mechanically, the same effective mass based Schrodinger equation is solved with the masses quoted from references [55, 56]. Due to their different perpendicular masses, the heavy holes form the first set of energy levels which are relatively low, and the light holes form the second set with higher confined energies. SCHRED V1.0 also has the capability of treating the electron/hole density in the inversion layer classically by using either Maxwell–Boltzmann or Fermi–Dirac statistics.

In doing bulk structure quantum mode simulation, SCHRED V1.0 can not only solve the effective mass based Schrödinger equation for inversion layer carriers, but also can solve the equation for accumulation layer carriers, for example, if the bulk is *p*-type silicon, in the inversion range, electrons are treated quantum mechanically, whereas in the accumulation range, holes are treated quantum mechanically. This is a feature that many other simulators do not offer.

In doing SOI quantum mode simulation, both electrons and holes are treated quantum mechanically at the same time. This is because in most cases, the SOI bodies are undoped or lightly doped, and the two dielectric gates confine the carriers in both inversion and accumulation regimes, therefore, the quantum effects can be equally important for both electrons and holes at low biases.

For both simulation modes, (classical or quantum mechanical) if the gate contacts are polysilicon, the charge density on the gates will always be computed classically. The gate dielectric constant can be specified different from SiO_2. The latest version also allows different dielectrics for the top and bottom gates in a SOI structure. This eases the simulations of effects of exotic insulator materials on device performance. Typical outputs of the solver are the spatial variations of the conduction-band edge and 3D charge density in the body; 2D surface charge density, average distance of the carriers from the interface; inversion layer capacitance C_{inv}, depletion layer capacitance C_{depl}, total gate capacitance C_{tot} and in the case of capacitors with poly-silicon gates, it also calculates the poly-gate capacitance C_{poly}. When choosing quantum-mechanical description of the electron density in the channel, it also provides the subband energies, the subband population, and the wavefunction variations in the body.

Schred is written in Fortran 77. The program is more efficient compared to other 1D Schrödinger–Poisson self-consistent simulators. A simplified flow-chart of the SCHRED V1.0 code is given in Fig. 2.17.

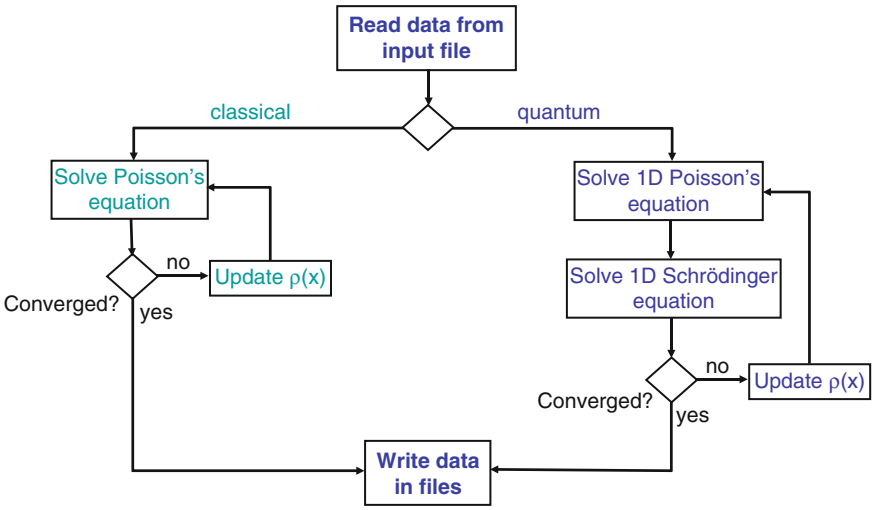


Fig. 2.17 Flow-chart of Schred V1.0

Examples of the application of SCHRED V1.0 can be found in [57–59] and in the sub-sections below.

Representative Simulation Results Obtained with SCHRED V1.0

Example 1: Semiclassical Versus Quantum Behavior

A first set of important simulation results that can be obtained with SCHRED V1.0 is the comparison between the semi-classical and quantum-mechanical models and how that affects the shape of the electron density and the magnitude of the sheet charge density. For that purpose we simulate an MOS capacitor with oxide thickness $t_{ox} = 1\,\text{nm}$, substrate doping $N_A = 10^{18}\,\text{cm}^{-3}$ and applied gate bias of 1 V. The metal workfunction is assumed to be equal to the semiconductor affinity.

The simulation results for the sheet electron density obtained with SCHRED V1.0 are: $N_s(\text{semi} - \text{classical}) = 1.43 \times 10^{13}\,\text{cm}^{-2}$ and $N_s(\text{quantum}) = 1.08 \times 10^{13}\,\text{cm}^{-2}$. These results indicate that the semiclassically calculated sheet electron density is about 30% higher than the quantum-mechanically calculated sheet electron density. There are two reasons for this: (1) the bandgap widening effect in the case of the quantum-mechanical model due to the shift of the first allowed state in the conduction band by 200.47 meV, and (2) the charge set-back from the interface because the wavefunction vanishes right at the interface, which leads to effective oxide thickness larger than the physical oxide thickness, thus leading to transconductance degradation. The charge set-back is clearly seen from the results shown in Fig. 2.18 where we plot the semi-classically calculated total electron density and the quantum-mechanically calculated total electron density. We see that

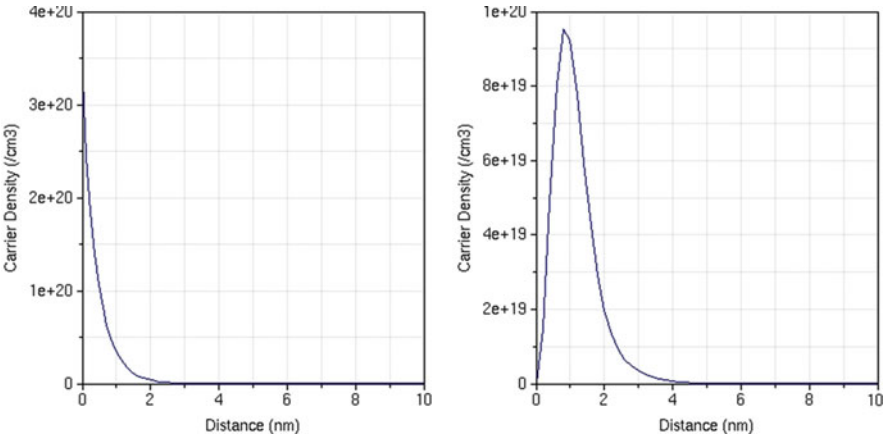


Fig. 2.18 Semiclassical *(left panel)* and quantum-mechanical *(right panel)* electron density

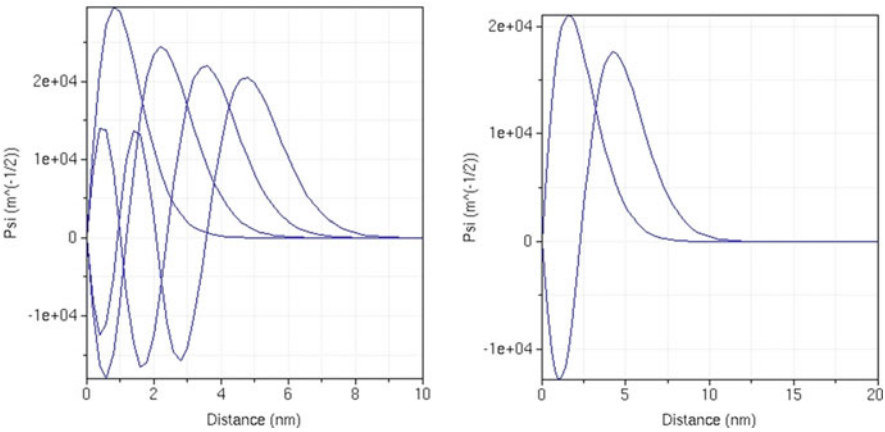


Fig. 2.19 Wavefunctions of the unprimed *(left)* and primed *(right)* ladder of subbands

the semiclassical charge density peaks at the interface as it is exponentially dependent of the negative of the potential, whereas the quantum-mechanically calculated electron density is zero at the interface and peaks at few angstroms away from the interface.

For the case of the quantum-mechanical model we have taken 4 subbands from the unprimed ladder of subbands and 2 subbands from the primed ladder of subbands. The spatial variation of the corresponding wavefunctions is shown in Fig. 2.19. There are several important things that can be observed from the results shown in Fig. 2.19. First, the shape of the wavefunctions resembles Airy functions that are solution to the 1D Schrödinger equation with linear potential energy term. Second, if we compare the first two wavefunctions from both the unprimed and

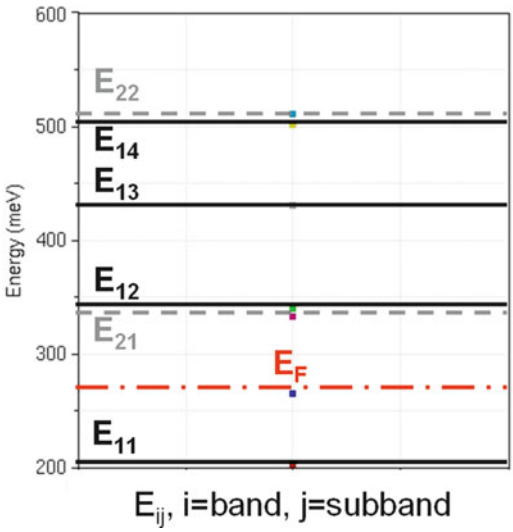


Fig. 2.20 Energy levels values from the unprimed and primed ladder of subbands

primed latter of subbands, then we see that the unprimed wavefunctions are more squeezed as the energies are lower and for those energies (see Fig. 2.20) the well is squeezed, therefore there exists larger localization of the carriers. Third, the first wavefunction has zero intersections with the x-axis, the second one has one, the third one has two, etc.

The corresponding energy levels of the unprimed and primed ladder of subbands are shown in Fig. 2.20. We see that the Fermi-level is above the first subband, therefore the semiconductor is degenerate. More importantly, we see that as we go higher in energy, the well widens and the energy level separation becomes smaller and smaller.

Example 2: Total Capacitance Degradation for Old and New Technology Nodes

In this second example we examine degradation of the total gate capacitance as a function of technology node. We consider what we call state of the art device technology, which is essentially the MOS capacitor discussed in Sect. 2.2.1. Regarding the older device technology MOS capacitor, its parameters are as follows: $N_A = 10^{16}\,\mathrm{cm}^{-3}$ and $t_{ox} = 40\,\mathrm{nm}$. The results of the simulations are presented in Figs. 2.21 and 2.22. There are several noteworthy features that can be deduced from the results shown.

For the case of state-of-the-art MOS capacitors, looking at the capacitances obtained for the case when the electron density is treated classically and quantum-mechanically, we observe two very important things: (1) there is a threshold voltage shift due to the quantum-mechanical size-quantization effect, and (2) there is a significant degradation of the total gate capacitance when using the quantum charge model that effectively degrades the device transconductance. The total capacitance degradation can be explained by examining the results for the average distance of the

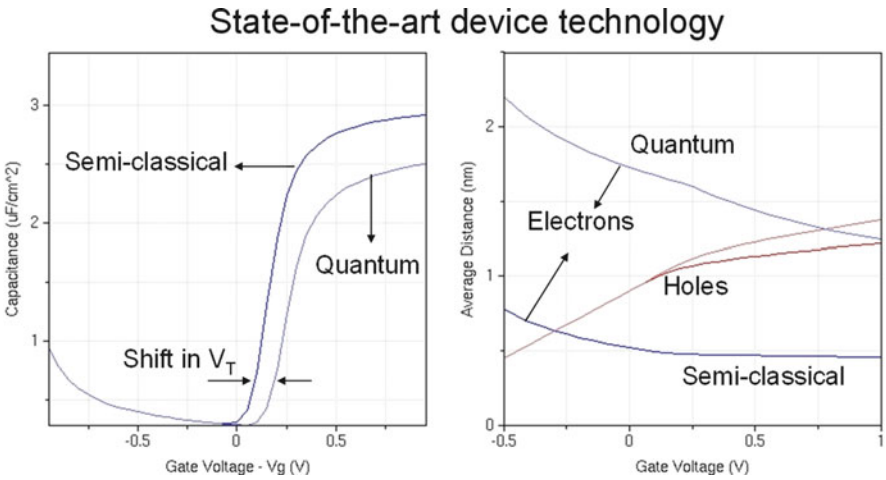


Fig. 2.21 *Left panel* – Total gate capacitance vs. gate voltage for state of the art device technology. *Right panel* – Average distance of the carriers from the interface

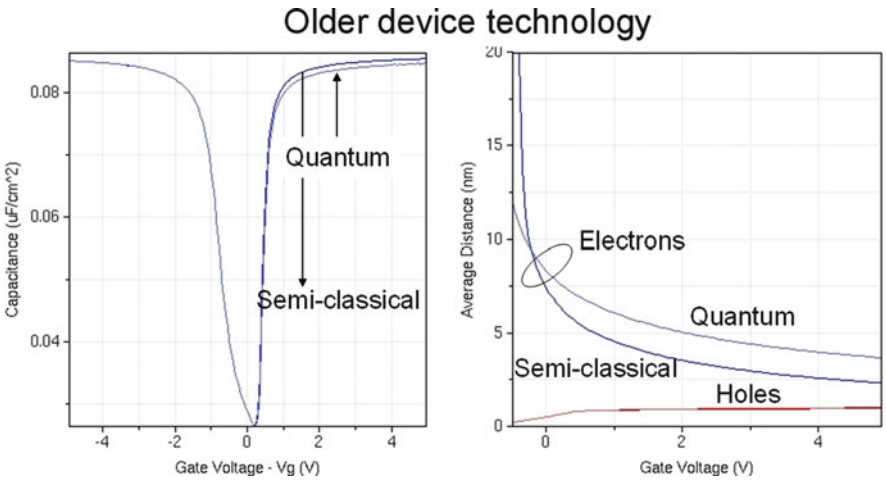


Fig. 2.22 *Left panel* – Total gate capacitance vs. gate voltage for older device technology. *Right panel* – Average distance of the carriers from the interface

electrons from the interface (Fig. 2.21 – Right panel). We see that classically carriers are about three times closer to the semiconductor/oxide interface when compared to the quantum case. The average distance in a way is a measure of the effective oxide thickness and quantum charge model leads to larger effective oxide thickness; therefore smaller transconductance.

For the case of older technology devices, looking at the results for the total gate capacitance shown in the left panel of Fig. 2.22, we might safely say that quantum

effects are not important as the total capacitance degradation is negligible. This can be attributed to the lower energy levels due to the wider well because of two orders of magnitude lower doping. As the well is wider, the average distance of the electrons from the interface is larger but that does not lead to transconductance degradation because the oxide thickness is 40 nm (40 times larger than in state-of-the-art devices).

From these two examples we might conclude that when modeling novel technology devices, quantum effects must be accounted for to properly determine the threshold voltage and total gate capacitance.

Example 3: Single Versus Dual Gate Capacitors

One of the primary reasons for device degradation at shorter channel lengths in FD SOI devices is the encroachment of drain electric field in the channel region. As shown in Fig. 2.23, the gate electrode shields the channel region from those lines at the top of the device, but electric field lines penetrate the device laterally and from underneath, through the buried oxide and the silicon wafer substrate causing the undesirable DIBL for the charge carriers.

To prevent the encroachment of electric field lines from the drain on the channel region, special gate structures can be used as shown in Fig. 2.24. Such "multiple-gate" devices include double-gate transistors, triple-gate devices such

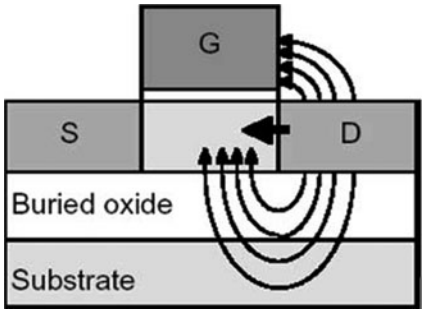


Fig. 2.23 Electric field lines from the drain

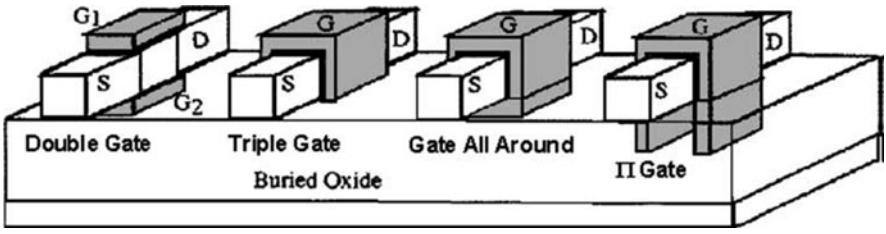


Fig. 2.24 Double-gate, triple-gate, gate all around (GAA), and Π-gate SOI MOSFETs

as the quantum wire [60], the FinFET [61] and Π-channel SOI MOSFET [62], and quadruple-gate devices such as the gate-all-around device [63], the DELTA transistor [64], and vertical pillar MOSFETs [65].

The double-gate device structure allows for termination of the drain electric field at the gates and leads to a more scalable FET. The double-gate concept was first reported in 1984 [66] and has been fabricated by several groups since then. The salient features of the DG FET (Fig. 2.24) are: (1) control of short-channel effects by device geometry, as compared to bulk FET, where the short-channel effects are controlled by doping (channel doping and/or halo doping); and (2) a thin silicon channel leading to tight coupling of the gate potential with the channel potential. These features provide potential DG FET advantages that include: (1) reduced 2D short-channel effects leading to a shorter allowable channel length compared to bulk FET; (2) a sharper subthreshold slope (60 mV/dec compared to 80 mV/dec for bulk FET) which allows for a larger gate overdrive for the same power supply and the same off-current; and (3) better carrier transport as the channel doping is reduced (in principle, the channel can be undoped). Reduction of channel doping also relieves a key scaling limitation due to the drain-to-body band-to-band tunneling leakage current. A further potential advantage is more current drive (or gate capacitance) per device area; however, this density improvement depends critically on the specific fabrication methods employed and is not intrinsic to the device structure. The most common mode of operation of the DG FET is to switch the two gates simultaneously.

In this exercise, we compare the performance of single-gate vs. double-gate MOSFET device structure by considering the double-gate option in SCHRED V1.0. We assume metal gates and the second gate is set to $V_{G2} = 1\,V$, and we sweep the first gate V_{G1}. The simulation results of the sheet electron density in the channel for single-gate and double-gate MOS capacitor are shown in Fig. 2.25. We use $t_{ox} = 1\,nm$ and $N_A = 10^{18}\,cm^{-3}$. For the double-gate MOS capacitor the body thickness is 10 nm. Evidently, we have almost twice the number of electrons in the channel region in the double-gate structure when compared to the single-gate structure.

Example 4: Dual Gate Capacitors – Volume Inversion

The thickness and/or width of multi-gate FETs are reaching values that are less than 10 nanometers. Under these conditions the electrons in the channel (if we take the example of an n-channel device) form either a two-Dimensional Electron Gas (2DEG) if we consider a double-gate device or a one-Dimensional Electron Gas (1DEG) if we consider a triple or quadruple-gate MOSFET. This confinement is at the origin of the "volume inversion" effect and yields an increase of threshold voltage when the width/thickness of the devices is reduced. The volume inversion effect is illustrated in Figs. 2.26 and 2.27, where we plot the electron density profile vs. gate voltage and the sheet electron density vs. body thickness, respectively.

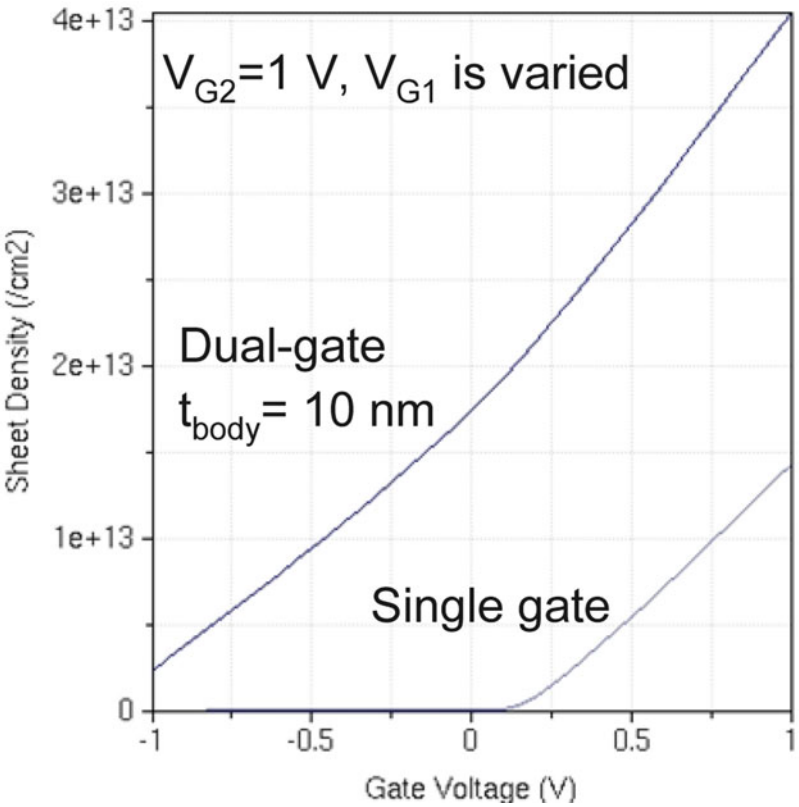


Fig. 2.25 Sheet electron density in a single-gate and double-gate structure as a function of the front gate voltage

2.2.2 SCHRED Second Generation Capabilities

Theoretical Model and Implementation Details

The theoretical model implemented is as follows. First user chooses one of the material systems described below. Then user specifies how many conduction bands are going to be taken into consideration. Then, for each specified conduction band (or pair of bands in the case of Si or strained-Si) the user specifies the effective masses. For the case of materials different than Si, the masses are taken to be isotropic. In the case of Si or strained-Si material system, the mass is assumed to be anisotropic, therefore crystallographic directions become important. Following the nomenclature of Rahman and co-workers [67], the user specifies the device, the crystal and the transport direction based on which one calculates the width, the confinement and the transport mass for each of the three pairs of ellipsoids of revolution

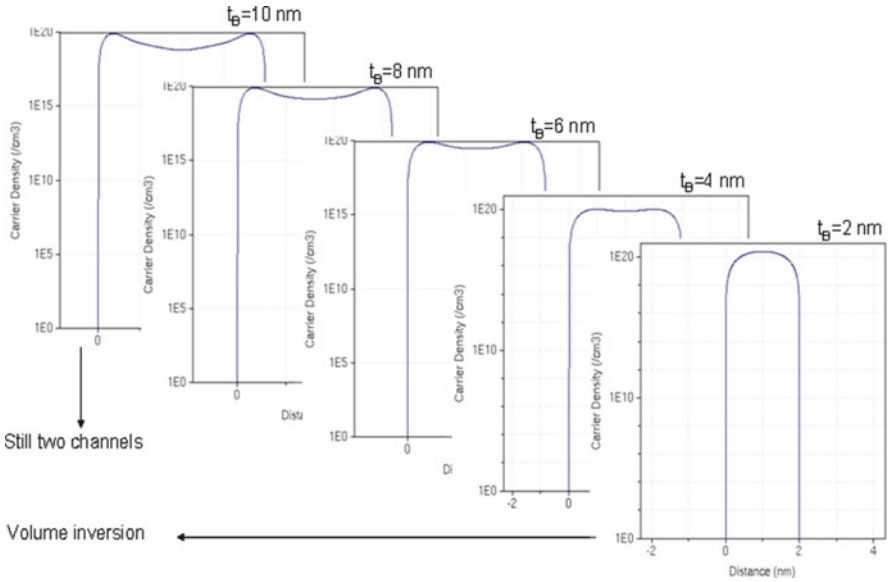


Fig. 2.26 Electron density profile for $V_{G1} = V_{G2} = 1V$

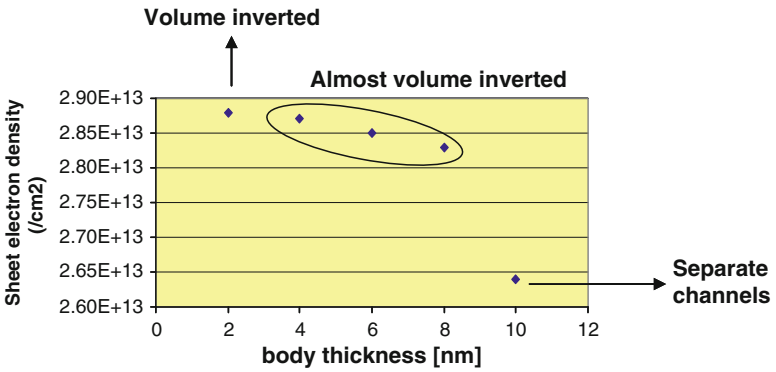


Fig. 2.27 Sheet electron density vs. silicon body thickness in the dual-gate structure

for the conduction band. Thus for a general conduction band ellipsoid (assuming 3 valleys) in the ellipse coordinate system (ECS),

$$E = \frac{\hbar^2 k_{||}^2}{2m_1} + \frac{\hbar^2 k_{\perp 1}^2}{2m_2} + \frac{\hbar^2 k_{\perp 2}^2}{2m_3} \tag{2.35}$$

For a given crystal coordinate system (CCS) and the ellipsoidal effective masses, we can write rotation matrix R_{E-C} for transforming components of an arbitrary

vector in CCS to its components in the ellipse co-ordinate system (ECS). Similarly we can write a rotation matrix R_{C-D} for transforming wave vector in the device co-ordinate system (DCS) to CCS. Thus we can write the inverse effective mass in the DCS as [68],

$$(M_D^{-1}) = R_{E\leftarrow D}^T (M_E^{-1}) R_{E\leftarrow D} \tag{2.36}$$

where

$$R_{E\leftarrow D} = R_{E\leftarrow C} R_{C\leftarrow D}, \tag{2.37}$$

and M_E^{-1} is a 3×3 diagonal matrix with m_l^{-1}, m_t^{-1}, m_t^{-1} along the diagonal. As a result, we can effectively model different orientations of Si or strained Si based on this approach for the effective mass calculation.

The valley offset in the conduction band in strained Si can be modeled using our three valley conduction band model. The various different effective masses for these three valleys can also be taken into consideration while solving the coupled system of Schrödinger–Poisson equations. The change in effective masses in the valence band of strained Si can also be included for the simulation.

As shown in Fig. 2.28, any material that can be expressed using a three valley conduction band system can be modeled by using our three valley conduction band model. This would enable us to model even those materials that are being researched at present. We can thus include in our simulation the different effective masses for the various conduction band and valence bands.

Because in some regimes of operation of the MOS capacitor there is no quantum-mechanical confinement and charge has to be treated classically, the effective density of states of the conduction band is calculated. Note that in SCHRED Second Generation holes at the moment are treated classically. In near future k.p method

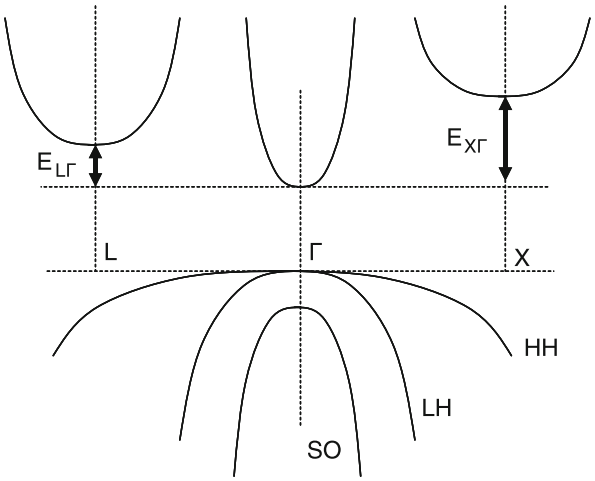


Fig. 2.28 General 3 valley conduction band model of a material

will be implemented to properly account for the warped valence bands and how they change under the influence of strain. User can choose whether to use semi-classical or quantum-mechanical charge description for the electrons. For the case of classical charge description the user has the option of Maxwell–Boltzmann and Fermi–Dirac statistics. The gate electrode can be treated as either a metal with user-defined workfunction or polysilicon. For simulations at low temperatures the users can also include partial ionization of the impurity atoms.

For the case of semiclassical charge description of the electrons and holes, only the linearized Poisson equation is solved using the LU decomposition method. When the electrons are treated quantum mechanically then a self-consistent solution of the 1D Poisson and the 1D Schrödinger equation is obtained. Note that the 1D Schrödinger equation is solved separately for each conduction band valley/valley pair. It is important to note that when finite difference approximation is applied to the 1D Schrodinger equation, a tri-diagonal non-symmetric coefficient matrix is obtained. Since the EISPACK routines that solve the eigenvalue problem are designed for symmetric coefficient tridiagonal matrices, a symmetrization procedure is necessary. This is achieved in the following manner. The discretized 1D Schrodinger equation is given by,

$$\sum_{j=1}^{n} A_{ij}\psi_j = \lambda\,\psi_i \tag{2.38}$$

where

$$A_{ij} = \begin{cases} -\dfrac{\hbar^2}{m^* x_i\,(x_i + x_{i-1})} & j = i+1 \\ \dfrac{\hbar^2}{m^* x_i\,(x_i + x_{i-1})} + \dfrac{\hbar^2}{m^* x_{i-1}\,(x_i + x_{i-1})} + V_i & j = i \\ -\dfrac{\hbar^2}{m^* x_{i-1}\,(x_i + x_{i-1})} & j = i-1 \\ 0 & otherwise \end{cases}$$

Thus, with the finite difference discretization of the 1D Schrödinger equation on a non-uniform mesh one arrives at a tridiagonal matrix that is not symmetric. The symmetrization of the coefficient matrix is achieved with the matrix transformation technique detailed below [69].

Let $x_i + x_{i-1}$ be L_i^2. Then, we have

$$A_{ij} = \begin{cases} -\dfrac{\hbar^2}{m^* x_i}\dfrac{1}{L_i^2} & j = i+1 \\ \left(\dfrac{\hbar^2}{m^* x_i} + \dfrac{\hbar^2}{m^* x_{i-1}}\right)\dfrac{1}{L_i^2} + V_i & j = i \\ -\dfrac{\hbar^2}{m^* x_{i-1}}\dfrac{1}{L_i^2} & j = i-1 \\ 0 & otherwise \end{cases}$$

Let $B_{ij} = L_i^2 A_{ij}$ or in matrix notation, $B = MA$, where M is the diagonal matrix with elements L_i^2, and B is tridiagonal and symmetric matrix. Thus the eigenvalue matrix (2.39) becomes,

$$B\psi = MA\psi = \lambda M\psi \tag{2.39}$$

The matrix M can be written as: $M = LL$, where L is a diagonal matrix with elements L_i. One can show that

$$L^{-1}BL^{-1}L\psi = L^{-1}LLA\psi = \lambda L^{-1}LL\psi, \tag{2.40}$$

or

$$H\varphi = \lambda\varphi, \tag{2.41}$$

where

$$H = L^{-1}BL{-}1, \tag{2.42}$$

and

$$\psi = L^{-1}\varphi. \tag{2.43}$$

Thus we can now solve using the symmetric matrix H, obtain the value of the φ matrix and from that obtain the value of ψ matrix – the eigenvectors.

Simulation Results

This section is divided into three parts. The first Sub-Section details the results from SCHRED Second Generation for the Silicon case. The following Sub-Section explains the results of SCHRED Second Generation in comparison with experimental results for a multi-valley semiconductor such as GaAs. The last Sub-Section compares experimental results of Strained Silicon for <100> transport orientation with the results of SCHRED Second Generation.

Example 1: Simulations of Regular Silicon for Specific Crystallographic Orientations

As shown in Table 2.1, the following orientations (wafer/transport/width directions) are simulated using SCHRED Second Generation.

We simulate MOS Capacitor with the following parameters: metal gate, substrate doping concentration of $10^{17}\,\mathrm{cm}^{-3}$, and oxide thickness of 4 nm. Two subbands are assumed for each of the three pairs of valleys. The resultant plots are then discussed. The effective masses for the different conduction band valley pairs are shown in Table 2.2 [67]. The mass m_z refers to the confinement effective mass and the mass m_{xy} refers to the product of the transport and width direction masses. This product contributes to the 2D density of states (DOS) mass.

Table 2.1 Different crystallographic orientations of silicon

(Wafer)/[Transport]/[Width]
(001) /[100]/[010]
$(111)/[\bar{2}11]/[0\bar{1}1]$
(110)/[001]/[00]

Table 2.2 Transport, width and confinement effective masses

Confinement direction	Transport, width and confinement effective mass	Valley 1	Valley 2	Valley 3
(001)	m_z	0.19	0.19	0.98
(110)	m_z	0.3189	0.3189	0.19
(111)	m_z	0.2598	0.2598	0.2598
(001)	m_{xy}	1.17	1.17	0.0361
(110)	m_{xy}	0.2223	0.2223	0.3724
(111)	m_{xy}	0.13604	0.13572	0.13572

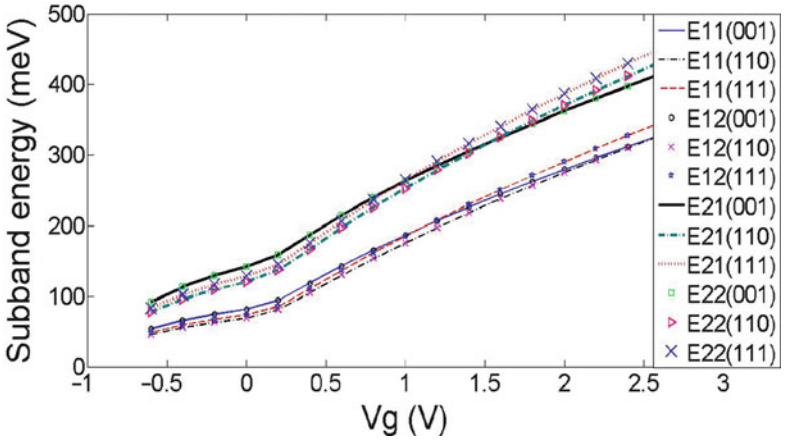


Fig. 2.29 Subband energy vs. applied voltage for valleys 1 and 2 (for various subband energy Eij, where i – denotes the subband, j – denotes the valley)

From the result shown in Fig. 2.29, it is evident that conduction band valley pair 1 has the lowest confinement mass for (001) confinement direction (see Table 2.2) and highest for (110) direction. Thus, the subband energies are lowest for the (110) direction and highest for the (001) direction. (The kinetic energy term in the Schrödinger equation will be the highest for the lowest mass, hence higher total subband energy). The valley pair 2 subband energies follow the same variation as the valley pair 1 subbands as they have the same set of masses in given directions and hence are equivalent to valley pair 1. The lower subband energies of valley pair 3 (unprimed set of subbands) as shown in Fig. 2.30, and are lower due to their higher confinement mass m_z (Table 2.2). As we increase the applied voltage, the potential well deepens, and the subband energies increase.

As shown in Fig. 2.31, the 2D sheet charge density is highest for the (001) orientation due to its lowest subband energy values. Thus we have lower sheet charge densities for the case of (110) which has higher subband energy than (001). In Fig. 2.32, the capacitance variation is presented for the three crystallographic directions. There is slight degradation for the total gate capacitance for orientations different than [001]. The most prominent result is shown in Fig. 2.33 where we

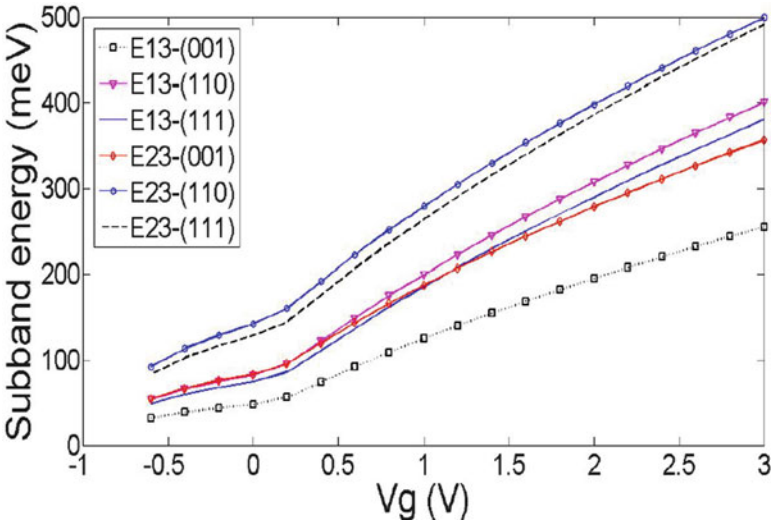


Fig. 2.30 Subband energy vs. applied voltage for valley3 (for various subband energy Eij, where i – denotes the subband, j – denotes the valley)

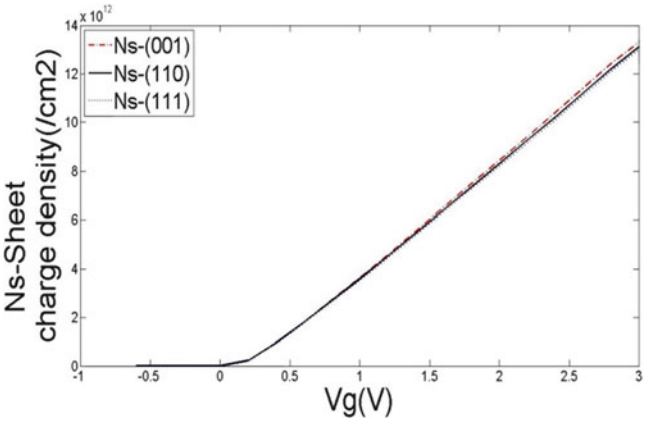


Fig. 2.31 Sheet charge density (Ns) vs. voltage

plot the average distance of the carriers from the interface as a function of the gate bias. We see that for [001] orientation we have the smallest average distance which means that in these devices interface roughness will play much higher role when compared to the other two crystallographic directions. This can significantly affect the on-current of the device fabricated in this material system.

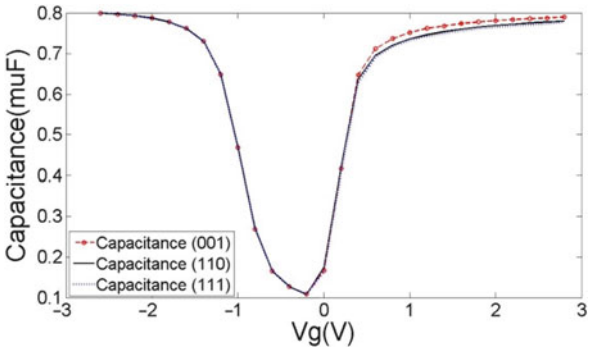


Fig. 2.32 Capacitance for the three confinement directions

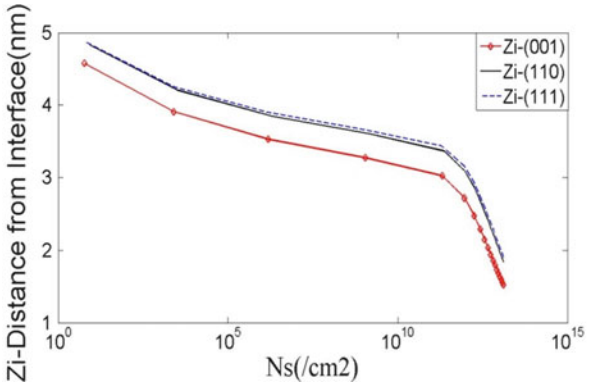


Fig. 2.33 Average distance of the carriers from the interface

Example 2: Gallium Arsenide MOS Capacitors

In order to verify the actual capability of SCHRED Second Generation in solving for multi-valley semiconductors, we had simulated MOS capacitors for a specific case of GaAs and compared our simulation results with the published data [70]. A substrate doping concentration of $10^{18}\,\mathrm{cm}^{-3}$ is used together with an oxide thickness of $t_{ox} = 16\,\mathrm{nm}$. The simulation runs have been performed for voltages in the range (-4 to $4\,\mathrm{V}$). We use three conduction band valleys (gamma, X and L valleys). We use two subbands for each of these valleys. The offsets between the valleys are included in the simulation.

From the results shown in Fig. 2.34 it can be seen that our results match much closer to the experimentally determined capacitance than the simulation results of [70]. The capacitance values match in the inversion and accumulation regions. We also observe that our results indicate a higher value of accumulation capacitance because we have not included hole confinement in the negative bias region of the simulation.

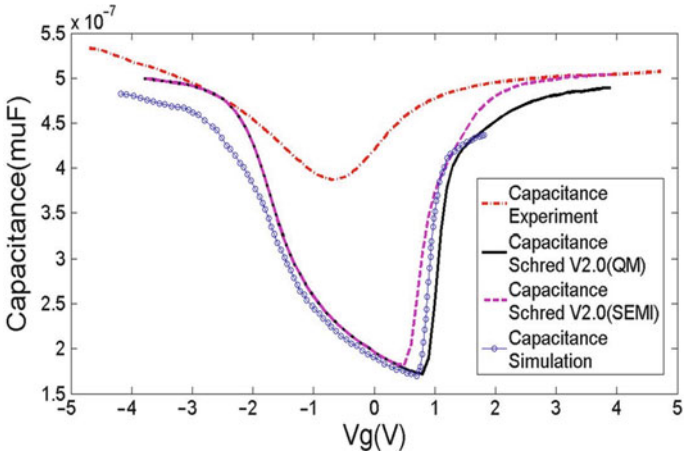


Fig. 2.34 GaAs capacitance for quantum mechanical (QM) and semi-classical case with experimental and simulation results from [70]

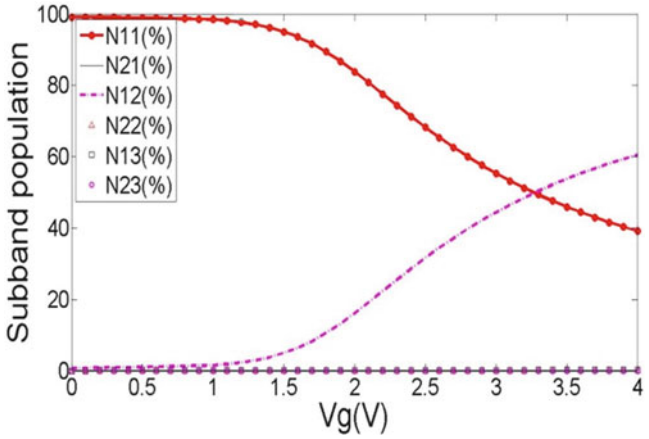


Fig. 2.35 Subband population

From the results presented in Figs. 2.35 and 2.36, we can clearly see that the subband population shifts from valley 1(gamma) to valley 2(L valley) as the gate voltage increases. More carriers are being excited to higher valleys, namely the L valley, as the applied voltage increases, thus increasing their population density. From the results presented in Fig. 2.36, it is also observed that only the lower subbands contribute to the majority of the population in a given valley, whereas the higher subbands are relatively unoccupied. This can be explained with the plot of the energy levels variation shown in Fig. 2.37.

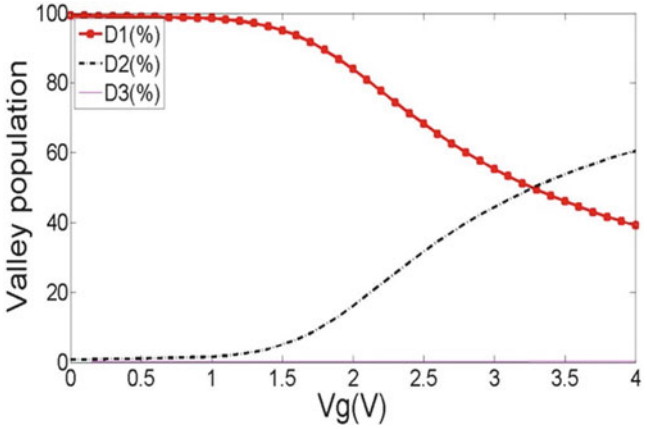


Fig. 2.36 Valley population

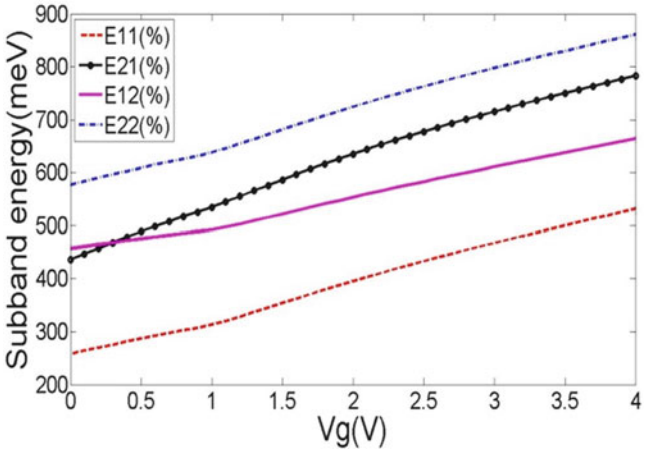


Fig. 2.37 Lowest two subband energies variation of the gamma and L valleys

Example 3: Strained Silicon

In the case of strained Si, strain on the Si material forces the valence bands degenerate levels to split; the heavy hole band crosses the light hole band and also the equi-energy Δ valleys are split into Δ_4 and Δ_2 conduction bands. This leads to change in the effective masses of the heavy hole and light hole valence bands (Figs. 2.38 and 2.39) and a change in the bandgap of the material.

Here we simulate to match experimental results of tensile strained Si (Silicon on silicon germanium). The experiment uses a polysilicon gate on a bi-axial strained Si layer on $Si_{0.8}Ge_{0.2}$. The experimental values are: polysilicon gate with doping concentration of $10^{20}\,cm^{-3}$ oxide thickness $t_{ox} = 1.33\,nm$, temperature $T = 300\,K$, substrate doping $N_A = 9 \times 10^{19}\,cm^{-3}$.

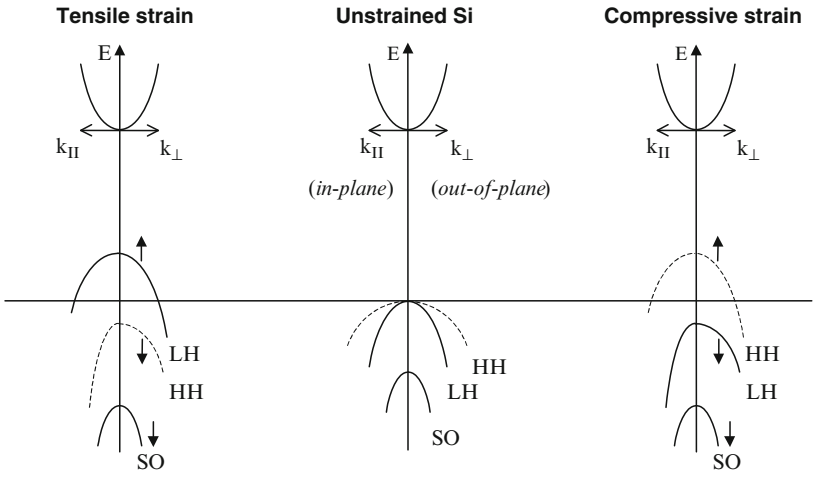


Fig. 2.38 Schematic band representation in strained layers under tensile and compressive strain, along with the unstrained case as a reference

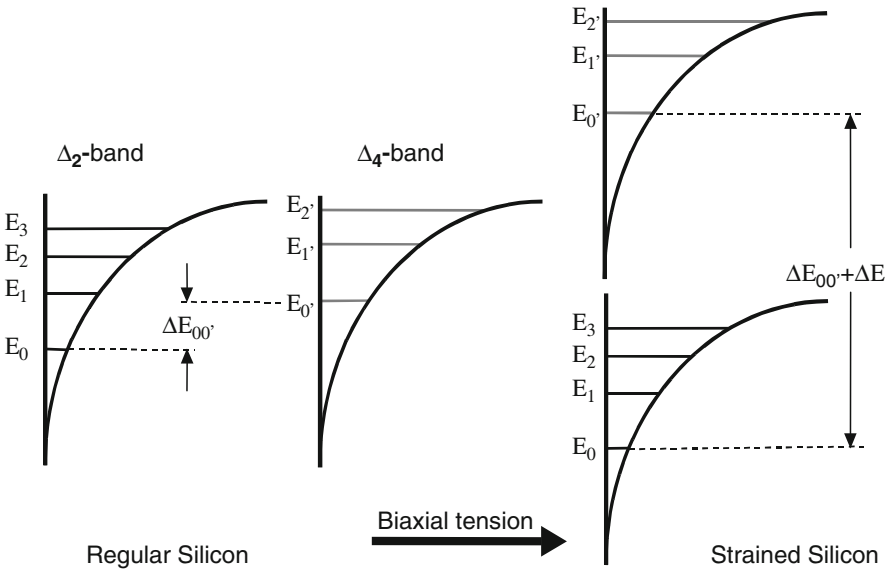


Fig. 2.39 Subband structure in the inversion layer of regular and surface-channel strained-Si layer

Our results in Fig. 2.40 closely match with the experimental results of [71]. The quantum capacitance matches with the experimental values in the inversion region, but differs in the accumulation and the depletion region due to the omission of the hole confinement in this work.

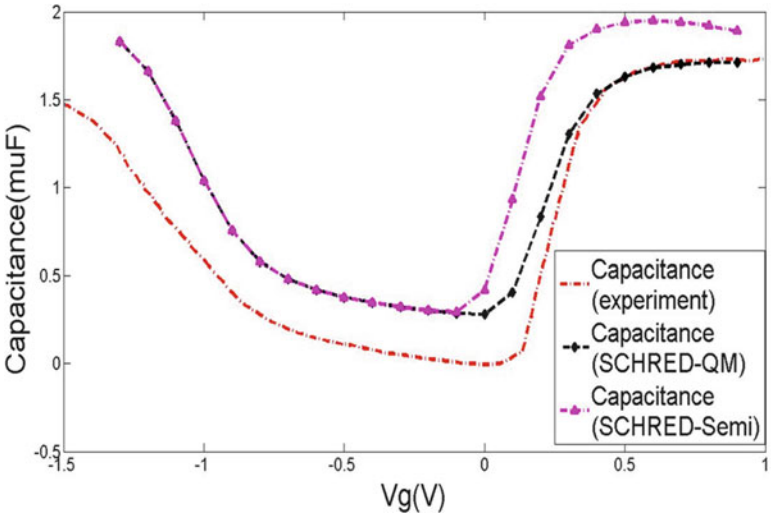


Fig. 2.40 Bi-axial strained on silicon (100) capacitance, experimental results from [71]

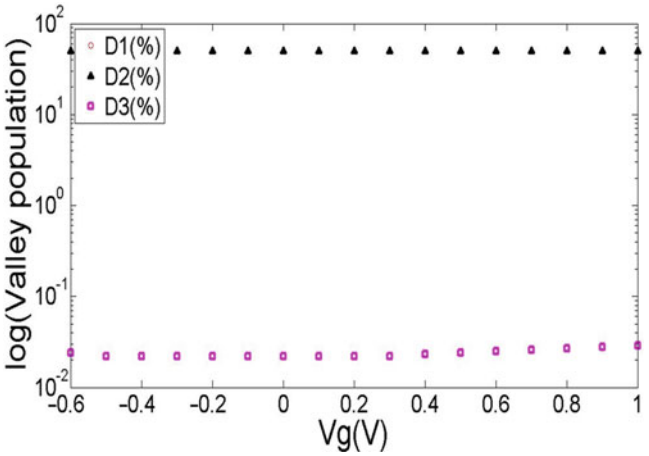


Fig. 2.41 Valley population

From the results presented in Fig. 2.41, we observe that, contrasting to the case of normal Si the population now shifts to the Δ_2 band(D2 valleys) from the Δ_4 band (D1 and D2 valleys) due to the application of the bi-axial strain (see Fig. 2.41), which makes the Δ_2 band to have a lower energy than the Δ_4 band.

Conclusions

This part of the research work presented in this book chapter has successfully created a nano-device simulator that can model MOS/SOS capacitors with the inclusion

of quantum effects, poly gate depletion, uniform/non-uniform doping, and user defined number of valleys, partial/complete ionization of carriers and several other features.

The simulator is built with a fast direct LU-decomposition Poisson solver that is coupled with the Schrödinger equation. The Schrödinger equation is solved in the bulk region using three point finite difference scheme, which results in a non-symmetric matrix (due to the non-uniform mesh used). This matrix is then transformed to a symmetric matrix using a matrix transformation technique. This transformed symmetric matrix is used to solve for eigenvalues and wavefunctions using the EISPACK routine.

2.3 Inclusion of Tunneling in Particle-Based Device Simulators

Tunneling is an important phenomenon in the operation of some devices in both the positive and the negative sense. For example, the negative differential characteristics in an Esaki diode (heavily doped p + /n+ junction – see Fig. 2.42) or in resonant tunneling diode are due to tunneling/resonant tunneling in these structures respectively. The peak to valley current is an important indicator on the quality of the device and larger the ratio, better is the device usability in oscillators.

Also, tunneling into the floating gate is necessary for the operation of EEPROM memories. Tunneling is the basic principle on which the operation of scanning tunneling microscopes is based, which revolutionized the understanding of surfaces and surface reconstructions in different semiconductor materials.

There are also instances in which tunneling is an undesired phenomenon, such as gate leakage in FET devices (see Fig. 2.43) or transistors with Schottky gate. In the case of FET devices, if the carriers tunnel through the tip of the barrier, then we call this tunneling process as Fowler–Nordheim tunneling. In small structures with thin oxides, carriers tunnel through the whole thickness of the oxide and in that case we have direct tunneling process.

The WKB (Wentzel, Kramers, Brillouin) approximation is a quasi-classical method for solving the one-dimensional (and effectively one-dimensional, such as radial) time-independent Schrödinger equation. The nontrivial step in the method is the connection formulas, that problem was first solved by Lord Rayleigh [72] and as Jeffries notes [73] "it has been rediscovered by several later writers" presumably referring to Wentzel, Kramers and Brillouin (WKB). A more accurate method for the calculation of the transmission coefficient in 1D tunneling structures is the transfer matrix approach which sometimes suffers from numerical overflow problems. To avoid these issues, a variant of this approach, the so-called scattering matrix approach is typically used. For 2D and 3D problems, the Usuki method [74] is the method of choice alongside with the Green's function approaches [75]. In what follows here, we first describe the WKB approximation on the example of tunneling through a triangular barrier, and then we discuss the transfer matrix approach on the example of a piecewise linear approximation of the potential barrier and its application in calculation of tunneling current in SOI Schottky MESFET.

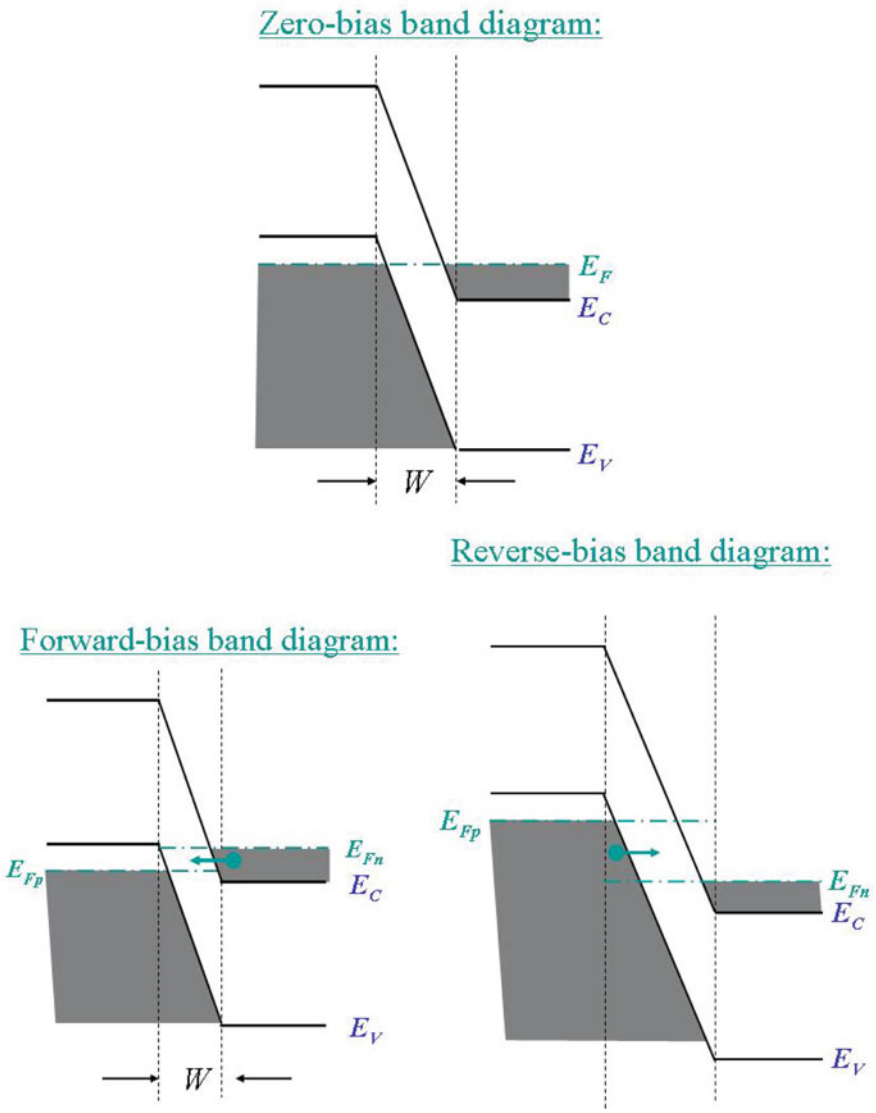


Fig. 2.42 Forward and reverse tunneling in heavily-doped PN (Esaki diodes). *Top panel* – Equilibrium band diagram, *bottom left panel* – forward bias conditions and *bottom right panel* – reverse bias conditions

2.3.1 WKB Approximation Used in Tunneling Coefficient Calculation

Consider a particle of mass m^* and energy $E > 0$ moving through some *slowly varying* potential $V(x)$. The particle's wave-function satisfies

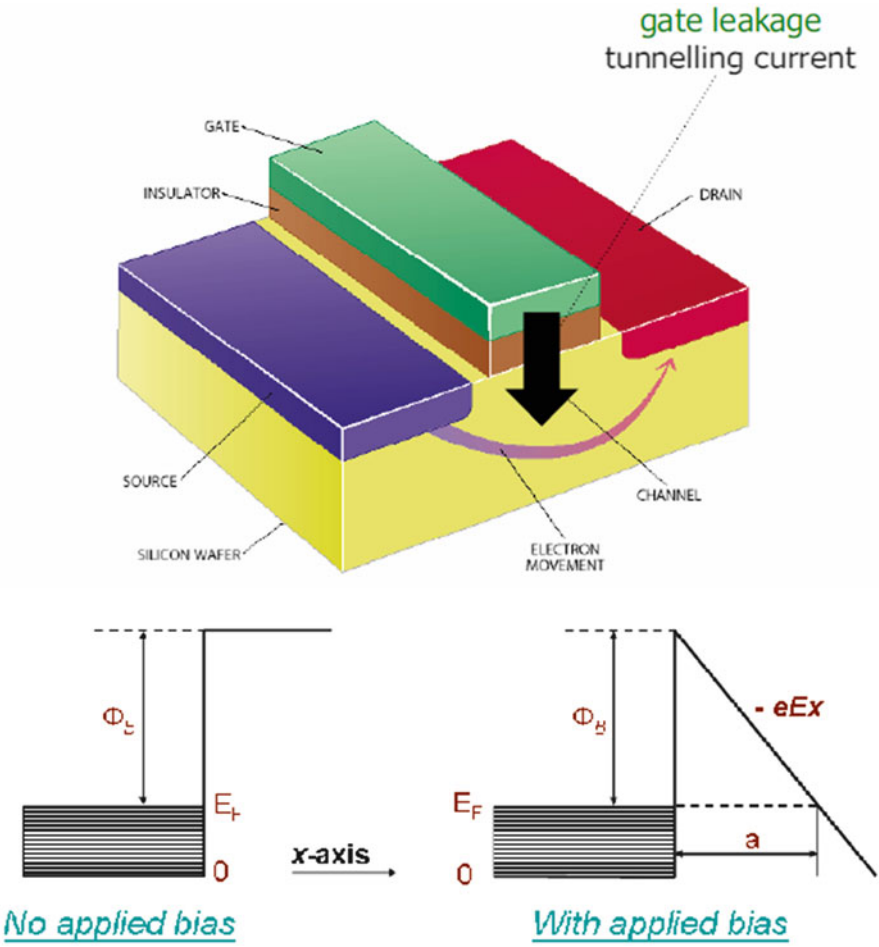


Fig. 2.43 Tunneling (gate leakage) limiting device miniaturization and leading to the introduction of gate stacks with high-k dielectrics (*top panel*). *Bottom panel* – Schematics of a tunnel barrier and the concept of Fowler–Nordheim tunneling

$$\frac{d^2\psi}{dx^2} = -k^2(x)\psi(x) \tag{2.44}$$

where

$$k^2(x) = \frac{2m^*[E - V(x)]}{\hbar^2} \tag{2.45}$$

Let us try a solution to (2.44) of the form

$$\psi(x) = \psi_0 \exp\left[\int_0^x ik(x')dx'\right] \tag{2.46}$$

where ψ_0 is a complex constant. Note that this solution represents a particle moving to the right with the continuously varying wavenumber $k(x)$. Substituting (2.46) into (2.44) gives

$$\frac{d^2\psi}{dx}^2 = ik'(x)\psi(x) - k^2(x)\psi(x) \tag{2.47}$$

where $k' = dk/dx$. From (2.44–2.47) it follows that (2.46) is a solution to (2.44) provided that the first term on its right-hand side is negligible compared to the second. This yields the validity criterion $|k'| \ll k^2$. In other words, the variation length-scale of $k(x)$ (which is approximately the same as the variation length-scale of $V(x)$) must be *much greater* than the particle's de Broglie wave-length (which is of order k^{-1}). Let us suppose that this is the case. Incidentally, the approximation involved in dropping the first term on the right-hand side of (2.47) is generally known as the *WKB approximation*. Similarly, (2.46) is termed a WKB solution. According to the WKB solution (2.46), the probability density remains constant: *i.e.*, $|\psi(x)|^2 = |\psi_0|^2$ as long as the particle moves through a region in which $E > V(\mathrm{x})$ and k(x) is consequently real (i.e., an allowed region according to classical physics).

Suppose, however, that the particle encounters a potential barrier (i.e., a region from which the particle is excluded according to classical physics). By definition, $E < V(x)$ inside such a barrier, and $k(x)$ is consequently imaginary. Let the barrier extend from $x = x_1$ to x_2, where $0 < x_1 < x_2$. The WKB solution inside the barrier is written

$$\psi(x) = \psi_1 \exp\left[-\int_{x_1}^{x} |k(x')|\, dx'\right] \tag{2.48}$$

where

$$\psi_1(x) = \psi_0 \exp\left[\int_0^{x_1} ik(x')dx'\right]. \tag{2.49}$$

Here, we have neglected the unphysical exponentially growing solution. According to the WKB solution, the probability density *decays exponentially* inside the barrier: i.e.,

$$|\psi(x)|^2 = |\psi_1|^2 \exp\left[-2\int_{x_1}^{x} |k(x')|\, dx'\right], \tag{2.50}$$

where $|\psi_1|^2$ is the probability density at the left-hand side of the barrier (i.e., $x = x_1$). It follows that the probability density at the right-hand side of the barrier (i.e., $x = x_2$) is

$$|\psi_2|^2 = |\psi_1|^2 \exp\left[-2\int_{x_1}^{x_2} |k(x')|\, dx'\right]. \tag{2.51}$$

Note that $|\psi_2|^2 < |\psi_1|^2$. Of course, in the region to the right of the barrier (i.e., $x > x_2$), the probability density takes the constant value $|\psi_2|^2$. We can interpret the

ratio of the probability densities to the right and to the left of the potential barrier as the probability $|T|^2$, that a particle incident from the left will tunnel through the barrier and emerge on the other side: i.e.,

$$T = \frac{|\psi_2|^2}{|\psi_1|^2} = \exp\left[-2\int_{x_1}^{x_2} |k(x')|dx'\right] \tag{2.52}$$

It is easily demonstrated that the probability of a particle incident from the right tunneling through the barrier is the same.

Note that the criterion for the validity of the WKB approximation implies that the above transmission probability is *very small*. Hence, the WKB approximation only applies to situations in which there is very little chance of a particle tunneling through the potential barrier in question. Unfortunately, the validity criterion breaks down completely at the edges of the barrier (i.e., at $x = x_1$ and x_2), since $k(x) = 0$ at these points. However, it can be demonstrated that the contribution of those regions, around $x = x_1$ and x_2, in which the WKB approximation breaks down to the integral in (2.52) is fairly negligible. Hence, the above expression for the tunneling probability is a reasonable approximation provided that the incident particle's de Broglie wave-length is much smaller than the spatial extent of the potential barrier.

Let us now apply the result given in (2.52) to the triangular barrier shown in Fig. 2.44. Upon the calculation of the integral in the exponent given by (2.52), one gets the transmission coefficient as,

$$T = \exp\left(-\frac{\pi m^{*1/2} E_G^{3/2}}{2\sqrt{2}\hbar e E}\right)\exp\left(-\frac{2E_z}{\bar{E}}\right), \tag{2.53}$$

where

$$\bar{E} = \frac{4\sqrt{2}\hbar e E}{3\pi\sqrt{m^* E_G}}, \tag{2.54}$$

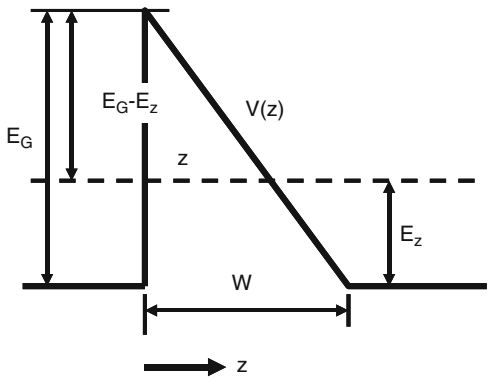


Fig. 2.44 Triangular potential barrier encountered by the electrons in an Esaki diode from Fig. 2.42 under forward and reverse bias conditions

and E is the electric field along the transport direction. The result given in (2.53) is then substituted in the Tsu–Esaki Formula for the current to get:

$$J_t = \frac{e^3 m^{*1/2} \xi V_a}{4\sqrt{2}\pi^2 \hbar^2 E_g^{1/2}} \exp\left(\frac{-4\sqrt{2m^*} E_G^{3/2}}{3e\hbar\xi}\right). \tag{2.55}$$

2.3.2 Transfer Matrix Approach for Piece-Wise Linear Approximation of the Potential Barrier

We next discuss the methodology for the calculation of the transmission probability and apply the technique for the calculation of the transmission coefficient through an arbitrary varying potential barrier. The exact method [76] that we use is based on the analytical solution of the Schrödinger equation across a linearly varying potential. In this case, the solution can be expressed as linear combination of Airy functions. Proper boundary conditions are imposed at the interface between adjacent linear intervals of the potential using a transfer matrix [77] procedure. The method for the calculation of the transmission coefficient is outlined below.

Let us consider a piecewise linear potential function such that the potential energy profile varies linearly in the region (a_{i-1}, a_i) (Fig. 2.45).

$$V(x) = V(a_{i-1}) + \frac{x - a_{i-1}}{a_i - a_{i-1}}[V(a_i) - V(a_{i-1})] = V_{i-1} + \frac{V_i - V_{i-1}}{a_i - a_{i-1}}(x - a_{i-1}) \tag{2.56}$$

The electric field profile is given by,

$$F_i = -\left.\frac{d\phi}{dx}\right|_i = \frac{1}{e}\left.\frac{dV}{dx}\right|_i = -\frac{V_i - V_{i-1}}{a_i - a_{i-1}}, \tag{2.57}$$

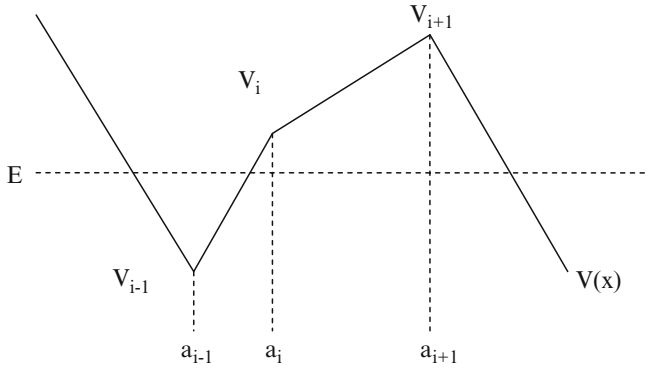


Fig. 2.45 Piecewise linear potential barrier

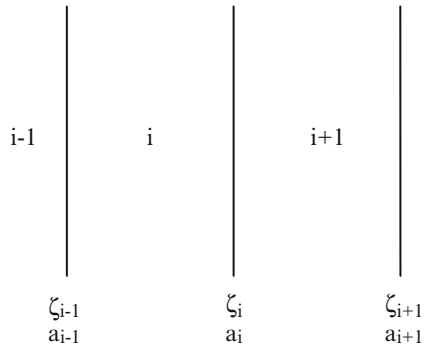


Fig. 2.46 Slicing of the region and corresponding variables in the slices

where V_i is in eV. Therefore,

$$V(x) = V_{i-1} + F_i(x - a_{i-1}) \tag{2.58}$$

Substituting back into the time-independent Schrödinger Wave Equation (TISE) gives (Fig. 2.46),

$$\begin{aligned}
&-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + V(x)\psi = E\psi, \\
&\Rightarrow -\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + [V_{i-1} + F_i(x - a_i)]\psi = E\psi, \\
&\Rightarrow -\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + F_i x\psi = (E + F_i a_i - V_{i-1})\psi, \\
&\Rightarrow -\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + F_i x\psi = \varepsilon'\psi.
\end{aligned} \tag{2.59}$$

We now define a dimensionless variable ξ such that

$$\xi = \left(\frac{2mF_i}{\hbar^2}\right)^{1/3} x - \frac{2m\varepsilon'}{\hbar^2}\left(\frac{\hbar^2}{2mqF_i}\right)^{2/3}. \tag{2.60}$$

Substituting (2.60) into (2.59) leads to

$$\frac{d^2\psi}{d\xi^2} - \xi\psi(\xi) = 0, \tag{2.61}$$

where $\varepsilon' = E + qF_i a_i - V_{i-1}$. The solutions of the reduced equation are the Airy functions and the modified Airy functions. Thus,

$$\psi_i = C_i^{(1)} A_i(\xi) + C_i^{(2)} B_i(\xi), \tag{2.62a}$$

and

$$\psi_{i+1}(\xi) = C_{i+1}^{(1)} A_i(\xi) + C_{i+1}^{(2)} B_i(\xi) \tag{2.62b}$$

From the continuity and the smoothness conditions for the wave function at $x = a_i$ we get

$$\psi_i(\xi_i) \;= \psi_{i+1}(\xi_i), \tag{2.63a}$$

$$\left.\frac{d\psi_i}{dx}\right|_{a_i} = \left.\frac{d\psi_{i+1}}{dx}\right|_{a_i} \Rightarrow \frac{d\psi_i}{dx} = \left.\frac{d\psi_i}{d\xi}\right|_{\xi_i} \frac{d\xi}{dx} = r_i \frac{d\psi_i}{d\xi}, \quad \left.\frac{d\psi_{i+1}}{dx}\right|_{a_i} = r_{i+1} \left.\frac{d\psi_{i+1}}{dx}\right|_{\xi_i} \tag{2.63b}$$

Therefore,

$$C_i^{(1)} A_i(\xi_i) + C_i^{(2)} B_i(\xi_i) \;= C_{i+1}^{(1)} A_i(\xi_i) + C_{i+1}^{(2)} B_i(\xi_i), \tag{2.64a}$$

$$r_i C_i^{(1)} A_i'(\xi_i) + r_i C_i^{(2)} B_i'(\xi_i) \;= r_{i+1} C_{i+1}^{(1)} A_i'(\xi_i) + r_{i+1} C_{i+1}^{(2)} B_i'(\xi_i). \tag{2.64b}$$

Rearranging (2.64a) and (2.64b) and writing them in a matrix form gives,

$$\begin{bmatrix} A_i(\xi_i) & B_i(\xi_i) \\ r_i A_i'(\xi_i) & r_i B_i'(\xi_i) \end{bmatrix} \begin{bmatrix} C_i^{(1)} \\ C_i^{(2)} \end{bmatrix} = \begin{bmatrix} A_i(\xi_i) & B_i(\xi_i) \\ r_{i+1} A_i'(\xi_i) & r_{i+1} B_i'(\xi_i) \end{bmatrix} \begin{bmatrix} C_{i+1}^{(1)} \\ C_{i+1}^{(2)} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} C_i^{(1)} \\ C_i^{(2)} \end{bmatrix} = M^{-1} \begin{bmatrix} A_i(\xi_i) & B_i(\xi_i) \\ r_{i+1} A_i'(\xi_i) & r_{i+1} B_i'(\xi_i) \end{bmatrix} \begin{bmatrix} C_{i+1}^{(1)} \\ C_{i+1}^{(2)} \end{bmatrix}$$

where

$$M^{-1} = \frac{1}{\det M} \begin{bmatrix} r_i B_i'(\xi_i) & -r_i A_i'(\xi_i) \\ -B_i(\xi_i) & A_i(\xi_i) \end{bmatrix}^T, \tag{2.65}$$

and $\det(M) = r_i [A_i(\xi_i) B_i'(\xi_i) - A_i'(\xi_i) B_i(\xi_i)] = \dfrac{r_i}{\pi}$. As a result of (2.65)

$$M^{-1} = \frac{\pi}{r_i} \begin{bmatrix} r_i B_i'(\xi_i) & -B_i(\xi_i) \\ -r_i A_i'(\xi_i) & A_i(\xi_i) \end{bmatrix},$$

and (2.65) becomes

$$\begin{bmatrix} C_i^{(1)} \\ C_i^{(2)} \end{bmatrix} = \frac{\pi}{r_i} \begin{bmatrix} r_i B_i'(\xi_i) & -B_i(\xi_i) \\ -r_i A_i'(\xi_i) & A_i(\xi_i) \end{bmatrix} \begin{bmatrix} A_i(\xi_i) & B_i(\xi_i) \\ r_{i+1} A_i'(\xi_i) & r_{i+1} B_i'(\xi_i) \end{bmatrix} \begin{bmatrix} C_{i+1}^{(1)} \\ C_{i+1}^{(2)} \end{bmatrix} = M_i \begin{bmatrix} C_{i+1}^{(1)} \\ C_{i+1}^{(2)} \end{bmatrix}. \tag{2.66}$$

Now let us consider the case for initial boundary between region 0 and region 1. In region 0 the wave function is described as plane wave and in region 1 it is a combination of Airy functions. Then

$$\psi_0 = C_0^{(1)} e^{ik_o x} + C_0^{(2)} e^{-ik_o x},$$
$$\psi_1(\xi) = C_1^{(1)} A_i(\xi) + C_1^{(2)} B_i(\xi). \tag{2.67}$$

The continuity of the wave function and of the derivative of the wave function leads to

$$C_0^{(1)} + C_0^{(2)} = C_1^{(1)} A_i(\xi_0) + C_1^{(2)} B_i(\xi_0),$$
$$ik_0[C_0^{(1)} - C_0^{(2)}] = r_1 C_1^{(1)} A_i'(\xi_0) + r_1 C_1^{(2)} B_i'(\xi_0). \tag{2.68}$$

Dividing the second equation by ik_o one gets

$$C_0^{(1)} - C_0^{(2)} = \frac{r_1}{ik_0} C_1^{(1)} A_i'(\xi_0) + \frac{r_1}{ik_0} C_1^{(2)} B_i'(\xi_0). \tag{2.69}$$

Then

$$2\,C_0^{(1)} = \left[A_i(\xi_0) + \frac{r_1}{ik_0} A_i'(\xi_0)\right] C_1^{(1)} + \left[B_i(\xi_0) + \frac{r_1}{ik_0} B_i'(\xi_0)\right] C_1^{(2)},$$
$$2\,C_0^{(2)} = \left[A_i(\xi_0) - \frac{r_1}{ik_0} A_i'(\xi_0)\right] C_1^{(1)} + \left[B_i(\xi_0) + \frac{r_1}{ik_0} B_i'(\xi_0)\right] C_1^{(2)}. \tag{2.70}$$

In summary,

$$\begin{bmatrix} C_0^{(1)} \\ C_0^{(2)} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}[A_i(\xi_0) + \frac{r_1}{ik_0} A_i'(\xi_0)] & \frac{1}{2}[B_i(\xi_0) + \frac{r_1}{ik_0} B_i'(\xi_0)] \\ \frac{1}{2}[A_i(\xi_0) - \frac{r_1}{ik_0} A_i'(\xi_0)] & \frac{1}{2}[B_i(\xi_0) + \frac{r_1}{ik_0} B_i'(\xi_0)] \end{bmatrix} \begin{bmatrix} C_1^{(1)} \\ C_1^{(2)} \end{bmatrix}. \tag{2.71}$$

We now consider the other boundary $[\mathrm{N},\ \mathrm{N}+1]$. In region N we have a combination of Airy functions and in region $\mathrm{N}+1$ we have plane waves. Hence, we have

$$\psi_N(\xi) = C_N^{(1)} A_i(\xi) + C_N^{(2)} B_i(\xi),$$
$$\psi_{N+1}(\xi) = C_{N+1}^{(1)} e^{ik_{N+1}x} + C_{N+1}^{(2)} e^{-ik_{N+1}x}. \tag{2.72}$$

The continuity of the wave function and of the derivative of the wave function then implies

$$\begin{aligned} C_N^{(1)} A_i(\xi_N) + C_N^{(2)} B_i(\xi_N) &= C_{N+1}^{(1)} e^{ik_{N+1}a_{N+1}} + C_{N+1}^{(2)} e^{-ik_{N+1}a_{N+1}}, r_N C_N^{(1)} A_i'(\xi_N) \\ &\quad + r_N C_N^{(2)} B_i'(\xi_N) \\ &= ik_{N+1}[C_{N+1}^{(1)} e^{ik_{N+1}a_N} - C_{N+1}^{(1)} e^{-ik_{N+1}a_N}. \end{aligned} \tag{2.73}$$

In matrix form this can be represented as,

$$\begin{bmatrix} C_N^{(1)} \\ C_N^{(2)} \end{bmatrix} = \frac{\pi}{r_n} \begin{bmatrix} r_N B_i'(\xi_N) + ik_{N+1} B_i(\xi_N) & r_N B_i'(\xi_N) - ik_{N+1} B_i(\xi_N) \\ -r_N A_i'(\xi_N) + ik_{N+1} A_i(\xi_N) & -r_N A_i'(\xi_N) - ik_{N+1} A_i(\xi_N) \end{bmatrix} M_1 \begin{bmatrix} C_{N+1}^{(1)} \\ C_{N+1}^{(2)} \end{bmatrix}. \tag{2.74}$$

Now, combining (2.66), (2.71), and (2.74), one finally arrives at the total transmission matrix of the system,

$$M_T = M_{FI}M_1M_2\ldots\ldots\ldots M_{N-1}M_{BI}\begin{bmatrix} e^{ik_{N+1}a_N} & 0 \\ 0 & e^{-ik_{N+1}a_N} \end{bmatrix}$$
$$= \begin{bmatrix} m_{11}^T & m_{12}^T \\ m_{21}^T & m_{22}^T \end{bmatrix}\begin{bmatrix} e^{ik_{N+1}a_N} & 0 \\ 0 & e^{-ik_{N+1}a_N} \end{bmatrix}. \tag{2.75}$$

The transmission coefficient is then given by,

$$T = \frac{k_{N+1}}{k_0}\frac{1}{|m_{11}^T|^2}, \tag{2.76}$$

where m_{11}^T is the element of the matrix $M_T = M_{FI}M_1M_2\ldots\ldots\ldots M_{N-1}M_{BI}$ and the various matrices that appear in (2.75) are defined as follows:

$$M_{FI} = \begin{bmatrix} \frac{1}{2}[A_i(\xi_0) + \frac{r_1}{ik_0}A_i'(\xi_0)] & \frac{1}{2}[B_i(\xi_0) + \frac{r_1}{ik_0}B_i'(\xi_0)] \\ \frac{1}{2}[A_i(\xi_0) - \frac{r_1}{ik_0}A_i'(\xi_0)] & \frac{1}{2}[B_i(\xi_0) + \frac{r_1}{ik_0}B_i'(\xi_0)] \end{bmatrix},$$
$$M_{BI} = \frac{\pi}{r_n}\begin{bmatrix} r_NB_i'(\xi_N) + ik_{N+1}B_i(\xi_N) & r_NB_i'(\xi_N) - ik_{N+1}B_i(\xi_N) \\ -r_NA_i'(\xi_N) + ik_{N+1}A_i(\xi_N) & -r_NA_i'(\xi_N) - ik_{N+1}A_i(\xi_N) \end{bmatrix},$$
$$M_i = \frac{\pi}{r_i}\begin{bmatrix} r_iB_i'(\xi_i) & -B_i(\xi_i) \\ -r_iA_i'(\xi_i) & A_i(\xi_i) \end{bmatrix}\begin{bmatrix} A_i(\xi_i) & B_i(\xi_i) \\ r_{i+1}A_i'(\xi_i) & r_{i+1}B_i'(\xi_i) \end{bmatrix}. \tag{2.77}$$

In the actual implementation of the method outlined above in the simulation of devices with Schottky barriers, we are considering the electrons between the gate and the buried oxide layer (in the active region) and we calculate the potential profile along the thickness of the device by solving Poisson's equation. Then, applying the Airy function transfer matrix method, we calculate the transmission probability for each particle in the MESFET device. On the basis of particle's position we calculate its potential energy. Then, we compare each particle's energy with the corresponding grid point potential energy. Now, using random number generation method, we evaluate whether each particle is going to tunnel through the Schottky barrier or not. If the transmission probability is greater than the random number then tunneling occurs. Once the particle tunnels, we use a rejection technique to make it inactive for the next iterative steps. For each time increment, we count the number of particles that tunnel through the barrier. After reaching a steady state condition, we calculate the tunneling current from the number of tunneled particles. We apply the piece-wise linear transfer matrix technique in a nonlinear potential barrier as shown in Fig. 2.47 to calculate the transmission probability. Following the technique, we have obtained the transmission probability which is shown in Fig. 2.48. From Fig. 2.48 it is observed that our result is properly matched with calculation previously performed by Lui et al. [78].

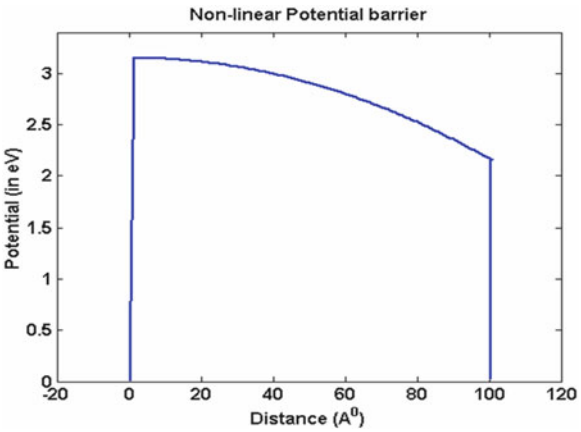


Fig. 2.47 Nonlinear potential barrier is used to calculate quantum mechanical transmission probability

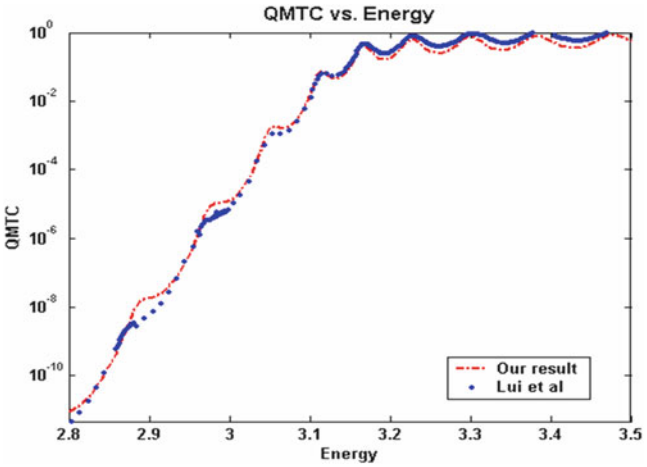


Fig. 2.48 Quantum mechanical transmission probability variation with respect to particle energy and validates our model's exactness

3 Discrete Impurity Effects

The pioneering experimental studies by Mizuno and co-workers [79] in the mid 1990s clearly demonstrated that threshold voltage fluctuations due to the discrete nature of the impurity atoms, are going to be a significant problem in future ultra-small devices. They had shown that the threshold voltage standard deviation is inversely proportional to the square root of the gate area, to the oxide thickness, and to the fourth root of the average doping in the device channel region. They also observed that the statistical variation of the channel dopant number accounts for

about 60% of the experimentally derived threshold voltage fluctuation. In a later study, Mizuno [80] also found that the lateral and vertical arrangement of ions produces variations in the threshold voltage that depend upon the drain and substrate biases. Horstmann and co-workers [81] investigated global and local matching of sub-100 nm n-channel metal-oxide-semiconductor (NMOS) and p-channel metal-oxide-semiconductor (PMOS)-transistors and confirmed the area law proposed in [80]. The empirical analytical expression by Mizuno was generalized by Stolk et al. [82] by taking into account the finite thickness of the inversion layer, the depth-distribution of the charge in the depletion layer and the influence of the source and drain impurity distributions.

Numerical drift-diffusion and hydrodynamic simulations [83–86] have also confirmed the existence of the fluctuations in the threshold voltage in ultra-small devices. Two-dimensional (2D) [87] and three-dimensional (3D) [88–91] ensemble Monte Carlo (EMC) particle-based simulations have also been carried out. An important observation was made in [10], where it was shown that there is a significant correlation between the threshold voltage shift and the actual position of the impurity atoms. A rather systematic analysis of the random dopant induced threshold voltage fluctuations in ultra-small metal-oxide-semiconductor field-effect transistors (MOSFETs) was carried out by Asenov [92] using 3D drift-diffusion device simulations and confirming previous results. Recent simulation experiments by Asenov and Saini [93] have shown that discrete impurity effects are significantly suppressed in MOSFETs with a δ-doped channel.

However, the majority of the above-mentioned simulation experiments, except [10,91], utilized 2D or 3D device simulators, in which the "discreteness" of the ions was only accounted for through the charge assignment to the mesh nodes. There, the long-range portion of the electron-ion forces are inherent in the mesh force and is found from the solution of the Poisson equation. The short-range portion of these interactions is either completely ignored or treated in the **k**-space portion of the EMC transport kernel (in particle based simulations) or via the doping dependence of the mobility (in drift-diffusion simulations). Because of the complexity and obscurity of the treatment of the Coulomb interaction in the MC simulations, a more direct approach has been introduced [10], in which the MC method is supplemented by a *molecular dynamics* (MD) routine. In this approach, the mutual Coulomb interaction among electrons and impurities is treated in the drift part of the MC transport kernel. Indeed, the various aspects associated with the Coulomb interaction, such as dynamical screening and multiple scatterings, are automatically taken into account. Very recently, the MC/MD method has been extended for spatially inhomogeneous systems. Since a part of the Coulomb interaction is already taken into account by the solution of the Poisson equation, the MD treatment of the Coulomb interaction is restricted only to the limited area near the charged particles. It is claimed that the full incorporation of the Coulomb interaction is indispensable to reproduce the correct electron mobility in highly doped silicon samples.

Although real space treatments eliminate the problem of double counting of the force, a drawback is that the 3D Poisson equation must be solved repeatedly to properly describe the self-consistent fields which consumes over 80% of the total

simulation time. To further speed up simulations, in this work a new idea has been proposed: to use a 3D Fast Multi-Pole Method (FMM) [94–97] instead. The FMM allows calculation of the field and the potential in a system of n particles connected by a central force within $O(n)$ operations given certain prescribed accuracy. The FMM is based on the idea of condensing the information of the potential generated by point sources in truncated series expansions. After calculating suitable expansions, the long range part of the potential is obtained by evaluating the truncated series at the point in question and the short range part is calculated by direct summation. The field due to the applied boundary biases is obtained at the beginning of the simulation by solving the Poisson equation. Hence the total field acting on each electron is the sum of this constant field and the contribution from the electron–electron and electron–impurity interactions handled by the FMM calculations. *The image charges, which arrise because of the dielectric discontinuity, are handled by the method of images*.

Quite recently, several groups, including ours [39], have shown that the Coulomb effects become even more prominent when the device size scales into the nm range. Even in undoped samples, a single unintentional dopant atom can cause significant fluctuations in the threshold voltage and therefore in the device on-state current due to the randomness of its position within the device active area. Thus, *proper inclusion of the short – range Coulomb interactions is a MUST when considering state of the art SOI FD-MOSFETs and alternate device structures, such as dual gate and FinFET devices*.

3.1 The P^3M *Method*

The particle-particle-particle-mesh (P^3M) algorithms are a class of hybrid algorithms developed by Hockney and Eastwood [98]. These algorithms enable correlated systems with long-range forces to be simulated for a large ensemble of particles. The essence of P^3M algorithms is to express the inter-particle force as a sum of a short-range part calculated by a direct particle–particle force summation and a long-range part approximated by the particle-mesh (PM) force calculation. Using the notation of Hockney, the total force on a particle i may be written as

$$F_i = \sum_{j \neq i} F_{ij}^{coul} + F_i^{ext}. \tag{2.78}$$

F_i^{ext} represents the external field or boundary effects of the global Poisson solution. F_{ij}^{coul}, is the force of particle j on particle i given by Coulomb's law as

$$F_{ij}^{coul} = \frac{q_i q_j}{4\pi\varepsilon} \frac{(r_i - r_j)}{|r_i - r_j|^3}, \tag{2.79}$$

where q_i and q_j are particle charges and r_i and r_j are particle positions. In a P^3M algorithm, the total force on particle i is split into two sums

$$F_i = \sum_{\substack{j \neq i \\ SRD}} F_{ij}^{sr} + \sum_{\substack{j \neq i \\ GD}} F_{ij}^{m}. \tag{2.80}$$

The first sum represents the direct forces of particles j on particle i within the short-range domain (SRD), while the second sum represents the mesh forces of particles j on particle i over the global problem domain (GD) that includes the effect of material boundaries and the boundary conditions on particle i. F_{ij}^{sr} is the short-range particle force of particle j on particle i, and F_{ij}^{m} is the long-range mesh force of particle j on particle i. The short-range Coulomb force can be further defined as,

$$F_{ij}^{sr} = F_{ij}^{coul} - R_{ij}, \tag{2.81}$$

where F_{ij}^{coul} is given by (2.79) and R_{ij} is called the reference force. The reference force in (2.81) is needed to avoid double counting of the short-range force due to the overlapping domains in (2.80). The reference force should correspond to the mesh force inside the short-range domain (SRD) and equal to the Coulomb force outside the short-range domain. In other words, a suitable form of the reference force for a Coulombic long-range force is one which follows the point particle force law beyond the cutoff radius r_{sr}, and goes smoothly to zero within that radius. Such smoothing procedure is equivalent to ascribing a finite size to the charged particle. As a result, a straightforward method of including smoothing is to ascribe some simple density profile $S(r)$ to the reference inter-particle force. Examples of shapes which are used in practice and give comparable total force accuracy are the uniformly charged sphere, the sphere with uniformly decreasing density

$$\boldsymbol{S}(\mathrm{r}) = \begin{cases} \dfrac{48}{\pi \boldsymbol{r}_{sr}^4}\left(\dfrac{\boldsymbol{r}_{sr}}{2} - \boldsymbol{r}\right), & \boldsymbol{r} \leq \boldsymbol{r}_{sr}/2 \\ 0, & \boldsymbol{otherwise,} \end{cases} \tag{2.82}$$

and the Gaussian distribution of density. The second scheme gives marginally better accuracies in 3D simulations. For this case the reference force can be obtained [99] as,

$$\begin{cases} \boldsymbol{R}_{ij}(\boldsymbol{r}) = \frac{qiqj}{4\pi\varepsilon} \times \frac{1}{35\boldsymbol{r}_{sr}^2}\left(224\xi - 224\xi^3 + 70\xi^4 + 48\xi^5 - 21\xi^6\right) & \xi = \frac{2r}{\boldsymbol{r}_{sr}} \text{ and } 0 \leq \boldsymbol{r} \leq \boldsymbol{r}_{sr}/2 \\ \boldsymbol{R}_{ij}(\boldsymbol{r}) = \frac{\boldsymbol{qiqj}}{4\pi\varepsilon} \times \frac{1}{35\boldsymbol{r}_{sr}^2}\left(\frac{12}{\xi^2} - 224 + 896\xi - 840\xi^2 - 224\xi^3 + 70\xi^4 + 48\xi^5 - 7\xi^6\right) & \boldsymbol{r}_{sr}/2 \leq \boldsymbol{r} \leq \boldsymbol{r}_{sr} \\ \boldsymbol{R}_{ij}(\boldsymbol{r}) = \frac{\boldsymbol{qiqj}}{4\pi\varepsilon} \times \frac{1}{\boldsymbol{r}^2} & \boldsymbol{r} > \boldsymbol{r}_{sr} \end{cases} \tag{2.83}$$

Hockney advocates pre-calculating the short-range force, $F_{ij}^{sr}(r)$ in (2.81) including the reference force above for a fixed mesh. It is important to extend the P^3M algorithm to nonuniform meshes for the purpose of semiconductor device simulation since practical device applications involve rapidly varying doping profiles and narrow conducting channels which need to be adequately resolved. Since the mesh force from the solution to the Poisson equation is a good approximation within about two mesh spaces, r_{sr} is locally chosen as the shortest distance which spans two mesh cells in each direction of every dimension of the mesh at charge i.

In order to incorporate the effects of material boundaries and boundary conditions, the reference force would be found most precisely in the short-range domain by associating particle j with the particle-mesh and calculating the resulting force on particle i with $F_i^{ext} = 0$. Since such a procedure would be required for each particle, it is obviously too costly for reasonable ensemble sizes and defeats the purpose of the P^3M algorithm [100]. Instead, it is desirable to use an approximation for this force, which minimizes the effects of the transition error in going from the long-range domain to the short-range domain. One approach developed in [100] is to choose a particular orientation of approaching particles relative to the mesh and find a radial approximation to the reference force. This method is straightforward and computationally efficient per particle for a fixed uniform mesh, but it is not easily adaptable to nonuniform meshes where the mesh force is not isotropic.

3.2 The Fast Multipole Method

FMM was first introduced by Rokhlin [95] and was later refined by Greengard [96] for the application of two and three-dimensional N-body problems whose interactions are Coulombic or gravitational in nature. In a system of N particles, the decay of the Coulombic or gravitational potential is sufficiently slow so that all interactions must be accounted for, resulting in CPU time requirements on the order of $O(N^2)$. On the other hand, the FMM requires an amount of work proportional to N to evaluate all interactions to within a round off error, making it practical for large-scale problems encountered in plasma physics, fluid dynamics, molecular dynamics, and celestial mechanics.

There have been a number of previous efforts aimed at reducing the computational complexity of the N-body problem. Assuming the potential satisfies Poisson's equation, a regular mesh is laid out over the computational domain and the method proceeds by: (1) interpolating the source density at mesh points; (2) using a fast Poisson solver to obtain potential values on the mesh; (3) computing the force from the potential and interpolating to the particle positions. The complexity of these methods is of the order of $O(N + M \log M)$, where M is the number of mesh points. The number of mesh points is usually chosen to be proportional to the number of particles, but with a small constant of proportionality so that $M \langle\langle N$. Therefore, although the asymptotic complexity for the method is $O(N \log N)$ the computational cost in practical calculations is usually observed to be proportional

to N. Unfortunately, the mesh provides limited resolution, and highly non-uniform source distributions cause a significant degradation of performance. Further errors are introduced in step (3) by the necessity for numerical differentiation to obtain the force. To improve the accuracy of particle-in-cell calculations, short-range interactions can be handled by direct computation, while far-field interactions are obtained from the mesh, giving rise to the so-called particle-particle-particle-mesh (P^3M) method described previously. While these algorithms still depend for their efficient performance on a reasonably uniform distribution of particles, in theory they do permit arbitrarily high accuracy to be obtained. As a rule, when the required precision is relatively low, and the particles are distributed more or less uniformly in a rectangular region, P^3M methods perform satisfactorily. However, when the required precision is high (for example in the modeling of highly correlated systems), the CPU time requirements of such algorithms tend to become excessive.

3.2.1 Multipole Moment

A multipole expansion is a series expansion which describes the effect produced by a given system in terms of an expansion parameter [95] that becomes smaller as the distance of the observation point from the source point increases. Therefore the leading order terms in a multipole expansion are generally the dominant. The first order behavior of the system at large distances can therefore be predicted from the first terms of the series, which is much easier to compute than the general solution.

Let r be the vector from the fixed reference point to a point in the system and r_1 be the vector from reference point to the observation point, and $d \equiv r_1 - r$ be the vector from a point in the system to the observation point. From the laws of cosines, d can be expressed as

$$d^2 = r_1^2 + r^2 - 2r_1 r\cos\varphi = r_1^2\left(1 + \frac{r^2}{r_1^2} - 2\frac{r}{r_1}\cos\varphi\right) \tag{2.84}$$

where $\cos\varphi \equiv \hat{r}.\hat{r_1}$. Therefore,

$$d = r_1\sqrt{1 + \frac{r^2}{r_1^2} - 2\frac{r}{r_1}\cos\varphi} \tag{2.85}$$

Let $\xi \equiv \frac{r}{r_1}$ and $y = \cos\varphi$. Then

$$\frac{1}{d} = \frac{1}{r_1}\left(1 - 2\xi y + \xi^2\right)^{-1/2} \tag{2.86}$$

But $(1 - 2\xi y + \xi^2)^{-1/2}$ is the generating function for Legendre Polynomials, i.e.

$$(1 - 2\xi y + \xi^2)^{-1/2} = \sum_{i=0}^{\infty} \xi^i P_i(y) \tag{2.87}$$

so,

$$\frac{1}{d} = \frac{1}{r_1} \sum_{i=0}^{\infty} \left(\frac{r}{r_1} \right)^i P_i(\cos\varphi) = \sum_{i=0}^{\infty} \frac{1}{r_1^{i+1}} r^i P_i(\cos\varphi). \tag{2.88}$$

Any physical potential that obeys a $1/d$ law can therefore be expressed as a multipole expansion,

$$V = \sum_{i=0}^{\infty} \frac{1}{r_1^{i+1}} \int r^i P_i(\cos\varphi)\, \rho(r) d^3 r. \tag{2.89}$$

In MKS unit,

$$V = \frac{1}{4\pi\varepsilon_0\varepsilon_r} \sum_{i=0}^{\infty} \frac{1}{r_1^{i+1}} \int r^i P_i(\cos\varphi) \rho(r) d^3 r, \tag{2.90}$$

where ε_0 is the permittivity of the free space, ε_r is the dielectric constant of the medium and $\rho(r)$ is the charge density.

3.2.2 How FMM Speeds Up the Computation?

In FMM *multipole moments* are used to represent distant particle groups and a *local expansion* is used to evaluate the contribution from distant particles in the form of a series. The multipole moment associated with a distant group can be *translated* into the coefficient of the local expansion associated with a local group. In FMM the computational domain is decomposed in a hierarchical manner with a quad-tree in two dimensions and an oct-tree in three dimensions to carry out efficient and systematic grouping of particles with tree structures. The hierarchical decomposition is used to cluster particles at various spatial lengths and compute interactions with other clusters that are sufficiently far away by means of the series expansions.

For a given input configuration of particles, the sequential FMM first decomposes the data-space in a hierarchy of blocks and computes local neighborhoods and *interaction-lists* involved in subsequent computations. Then, it performs two passes on the decomposition tree. The first pass starts at the leaves of the tree, computing *multipole expansion coefficients* for the Columbic field. It proceeds towards the root accumulating the multipole coefficients at intermediate tree-nodes. When the root is reached, the second pass starts. It moves towards the leaves of the tree, *exchanging* data between blocks belonging to the neighborhoods and interaction-lists calculated at tree-construction. At the end of the downward pass all long-range interactions have been computed. Subsequently, nearest-neighbor computations are performed directly to take into consideration interactions from nearby bodies. Finally, short- and long-range interactions are accumulated and the total forces exerted upon particles are computed. The algorithm repeats the above steps and simulates the evolution of the particle system for each successive time-step.

3.3 The Role of Discrete Impurities as Observed by Simulations and with Comparisons to Experiments

In the three subsequent subsections first the role of discrete impurities on the operation of conventional device designs is discussed, then unintentional dopants are being examined and finally the role of unintentional dopants on the FinFET transfer and output characteristics is being examined.

3.3.1 Previous Knowledge on Threshold Voltage and On-State Current Fluctuations in Sub-Micrometer MOSFET Devices

As already discussed in the introduction part of this book chapter, continued scaling of devices has led to a number of undesirable effects, including fluctuations in the threshold voltage that arise because of the discrete, or atomistic nature of the impurity atoms in the device active region. For better insight of the importance of this issue, we have considered a prototypical MOSFET with 0.07 μm channel length, 0.07 μm channel width and channel doping of $10^{18}\,\mathrm{cm}^{-3}$. The number of dopant atoms in the depletion region of this device is on the order of several hundreds, and well below 100 in the active region. In addition, there are regions where the impurity atoms cluster and other regions in which the impurity density is well below the average value expected from the doping level. With such a small number of the impurity atoms in the device active region, the local variations in the "doping concentration" across the channel become a significant factor in determining the threshold voltage, mobility and drain current characteristics. This in turn, causes considerable problems for circuit design, especially for circuits in which the devices must be well matched, such as operational amplifiers [101] and static random access memories [102]. The *SIA* roadmap technology requirements state that the variation in gate length should be less than 10% and the variation in threshold voltage should be less than 40 mV for devices in the 150 nm generation and beyond [103].

It is interesting to note that the existence of these surface potential fluctuations in MOS devices was postulated by Nicollian and Goetzberger [104] in order to explain the departures from the theoretical predictions in conductance vs. frequency measurements in MOS structures. In addition to their effect on the *ac*-conductance results, surface potential fluctuations were also found to have significant influence on a variety of other device characteristics, such as threshold voltage, transconductance, substrate current and off-state leakage currents. Experimental studies by Mizuno, Okamura, and Toriumi [6] have shown that the threshold voltage standard deviation is related to the average number of ionized impurities beneath the channel according to

$$\sigma_{vt} = \left(\frac{\sqrt[4]{q^3 \varepsilon_s \phi_b}}{\sqrt{2}\varepsilon_{ox}} \right) \frac{T_{ox} \sqrt[4]{N}}{\sqrt{L_{eff} W_{eff}}}, \tag{2.91}$$

where N is the average channel doping density, ϕ_b is the built-in potential, T_{ox} is the oxide thickness, L_{eff} and W_{eff} are the effective channel length and width, and ε_s and ε_{ox} are the semiconductor and oxide permittivity, respectively. They found that the statistical variation of the channel dopant number accounts for about 60% of the experimentally derived threshold voltage fluctuations. In a later study, Mizuno [81] also found that the lateral and vertical arrangement of ions produces variations in the threshold voltage dependence upon the drain and substrate bias. Quite recently, Horstmann, Hilleringmann and Goser [105], who investigated the global and local matching of sub-100 nm NMOS- and PMOS-transistors, confirmed the law of area given in (2.91). Also, Stolk et al. [106] generalized the analytical result by Mizuno and his co-workers by taking into account the finite thickness of the inversion layer, depth-distribution of charges in the depletion layer and the influence of the source and drain dopant distributions and depletion regions. For a uniform channel dopant distribution, the analytical expression for the threshold voltage standard deviation given in [107] simplifies to

$$\sigma_{vt} = \left(\frac{\sqrt[4]{q^3 \varepsilon_s \phi_b}}{\sqrt{3}} \right) \left[\frac{k_b T}{q} \cdot \frac{1}{\sqrt{4 \varepsilon_s \phi_b N_a}} + \frac{T_{ox}}{\varepsilon_{ox}} \right] \frac{\sqrt[4]{N}}{\sqrt{L_{eff} W_{eff}}}. \tag{2.92}$$

In (2.92), the first term in the square brackets represents the surface potential fluctuations whereas the second term represents the fluctuations in the electric field.

The purpose of this section is twofold. First, we will clarify some issues related to the origin of the threshold voltage fluctuations in ultra-small devices. The second, and more important issue discussed here is how discrete impurities affect device high-field characteristics, such as carrier drift velocity and the on-state currents in conventional MOSFETs.

The Role of the Short-Range e–e and e–i Interactions

To be able to study the effect of the proper inclusion of the short-range Coulomb force to the mesh force, the energy and position of several electrons were monitored during a simulation run. The simulated device has channel length $L_G = 80$ nm, channel width $W_G = 80$ nm and oxide thickness $T_{ox} = 3$ nm. The lateral extension of the source and drain regions is 50 nm. The channel doping equals $3 \times 10^{18}\,\mathrm{cm}^{-3}$. The applied bias is $V_G = V_D = 1$ V. Only those electrons that entered the channel region from the source side were "tagged" and their energy and position was monitored and used in the average energy calculation. The average velocity and the average energy of the electrons that reach the drain end of the device is shown in Fig. 2.49. From the average velocity simulation results, it follows that the short-range electron–electron (*e–e*) and electron–ion (*e–i*) interaction terms damp the velocity overshoot effect, thus increasing the transit time of the carriers through the device, in turn reducing its cut-off frequency (Fig. 2.49a). It is also quite clear that when we use the mesh force only, i.e. we skip the molecular dynamics (MD) loop that allows us to correct for the short-range *e–e* and *e–i* interactions, those electrons that enter the drain end of the device from the channel never reach equilibrium (Fig. 2.49b). Their average

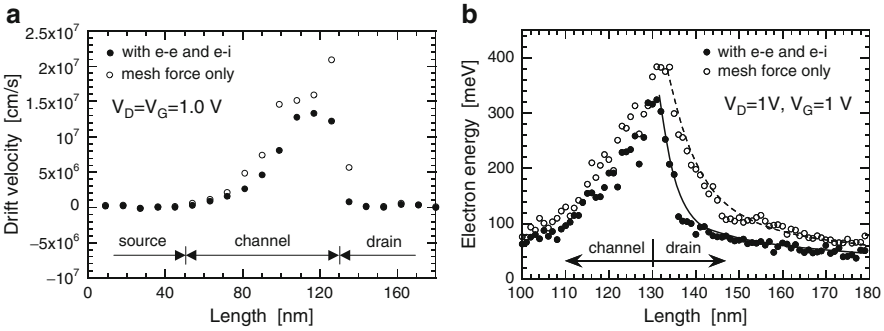


Fig. 2.49 (**a**) Average velocity of the electrons along the channel, with and without the inclusion of the *e–e* and *e–i* interactions. (**b**) Average energy of the electrons coming to the drain from the channel. The applied bias equals $VD = VG = 1\,\mathrm{V}$. *Filled (open) circles* correspond to the case when the short-range $e-e$ and $e-i$ interactions are included (omitted) in the simulations

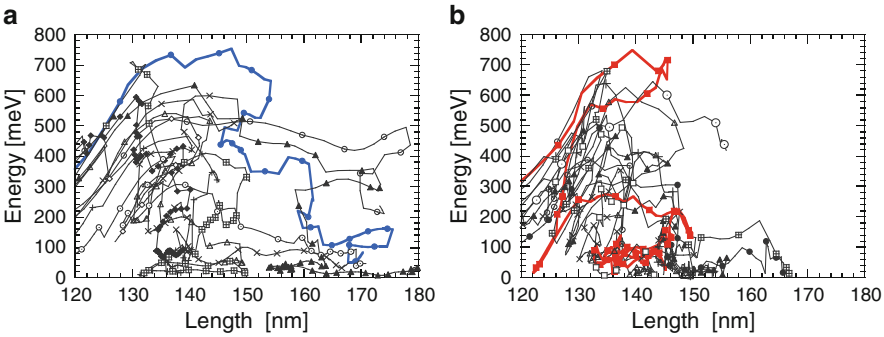


Fig. 2.50 (**a**) Phase-space trajectories of ten randomly chosen electrons for the case when the mesh force is only considered in the free-flight portion of the simulator. (**b**) Phase-space trajectories of ten randomly chosen electrons for the case when the short-range e–e and e–i interactions are included via our MD routine

energy is more than 60 meV far into the drain region. Also, the average energy peaks past the drain junction. The addition of the short-range Coulomb forces to the mesh force via the MD loop, leads to rapid thermalization of the carriers once they enter the drain region. The characteristic distance over which carriers thermalize is on the order of a few nm.

In Fig. 2.50, we show the phase-space trajectory of 10 randomly selected electrons that reach the drain region. We use $V_G = 0.5\,\mathrm{V}$, $V_D = 0.8\,\mathrm{V}$, $T_{ox} = 3\,\mathrm{nm}$, and $N_A = 3 \times 10^{17}\,\mathrm{cm}^{-3}$ in these simulations. Notice that some of the electrons reach the end of the device and are reflected back without losing much energy when we use the mesh force only (Fig. 2.50a). The addition of the short-range Coulomb force leads to very fast thermalization of the carrier energy once they enter the drain end (Fig. 2.50b). None of the randomly selected electrons reach the device boundary, as opposed to 3 out of 10 electrons reaching the boundary when the short-range Coulomb force is turned off.

Threshold Voltage Fluctuations

The threshold voltage fluctuations vs. device gate width, channel doping and oxide thickness, are shown in Fig. 2.51. Also shown in this figure are the analytical model predictions given by (2.91) and (2.92). The decrease of the threshold voltage

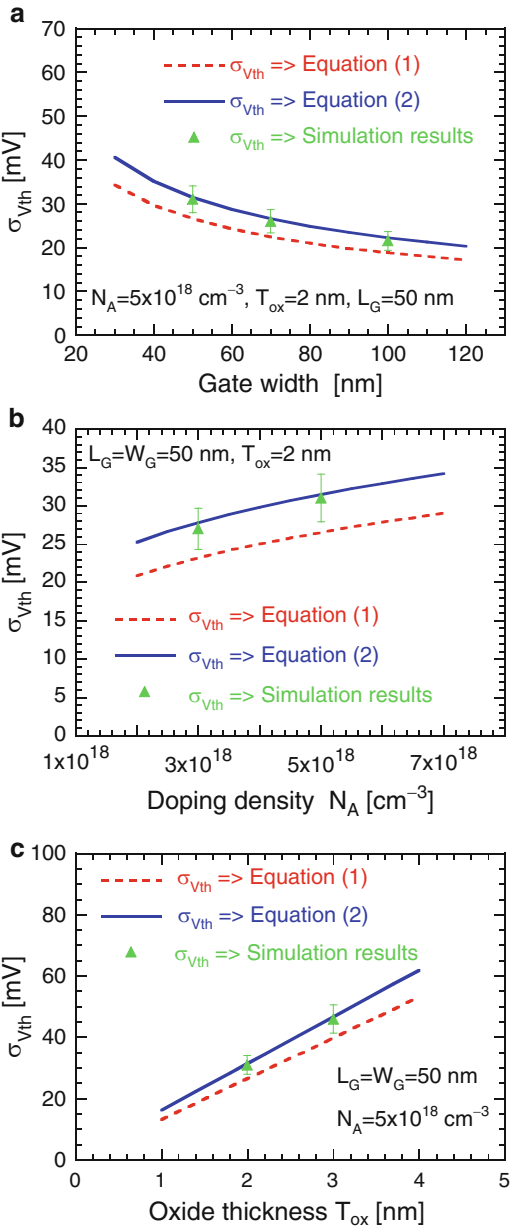


Fig. 2.51 Variation of the threshold voltage with (**a**) gate width, (**b**) channel doping, and (**c**) oxide thickness

fluctuations with increasing the width of the gate is due to the averaging effects, in agreement with the experimental findings by Horstmann et al. [82]. We want to point out that we still observed significant spread of the device transfer characteristics along the gate voltage axis even for devices with $W_G = 100\,\text{nm}$. This is due to the nonuniformity of the potential barrier, which allows for early turn-on of some parts of the channel. As expected, the increase in the channel doping leads to larger threshold voltage standard deviation $\sigma_{V_{TH}}$. These results also imply that the fluctuations in the threshold voltage can be even larger in devices in which counter ion implantation is used for threshold voltage adjustments. Similarly, the increase in the oxide thickness leads to linear increase in the threshold voltage standard deviation. The results shown in Fig. 2.51a–c also suggest that reconstruction of the established scaling laws is needed to reduce the fluctuations in the threshold voltage. In other words, within some new scaling methodology, T_{ox} should become much thinner, or N_A much lower that what the conventional scaling laws give.

Fluctuations in the On-State Currents

Besides investigating the threshold voltage fluctuations, our 3D EMC particle-based device simulator also allows us to investigate the fluctuations in the high-field characteristics, such as the saturation drain current. The variation of the drain current vs. the number of channel dopant atoms for the 15 devices from [107] described in terms of the number of dopants in Fig. 2.52a, is shown in Fig. 2.52c. Each device was simulated for a total of 4 ps. The gate voltage was set to 1.5 V and the drain voltage to 1.0 V. The drain current was measured by averaging the velocity of electrons in the channel over the last 2.4 ps of the simulation. It is important to note that at these bias conditions, the devices were in the saturation region of the $I_D - V_G$ curve, but were not velocity saturated.

As expected, as the number of channel dopant atoms increases, the drain current decreases due to the increase in the V_T. More importantly, for the five devices from the high-end of the distribution, due to the larger probability that some of the impurity atoms will be located near the semiconductor/oxide interface, there is larger fluctuation in the saturation current. This is also reflected in the average velocity of channel electrons vs. the number of dopant atoms in the channel, as shown in Fig. 2.52d. Again, the velocity decreases as the number of dopant atoms increases due to increased ionized impurity scattering. At the low end of the dopant number distribution, the average electron velocity is roughly the same for each dopant configuration. However, the fluctuation in the electron velocity increases with the number of dopant atoms, with a $3\times$ spread in the velocity seen for the devices at the high dopant number extreme.

The average electron velocity and device drain current characteristics were correlated to the number of dopant atoms in a 10 nm range at various depths. Figure 2.52 (Top right panel) shows a plot of the square of the correlation coefficient vs. depth (beneath the semiconductor/oxide interface). The correlation to the electron velocity is very high for the first 6 nm, and steadily decreases up to 18 nm depth, beyond

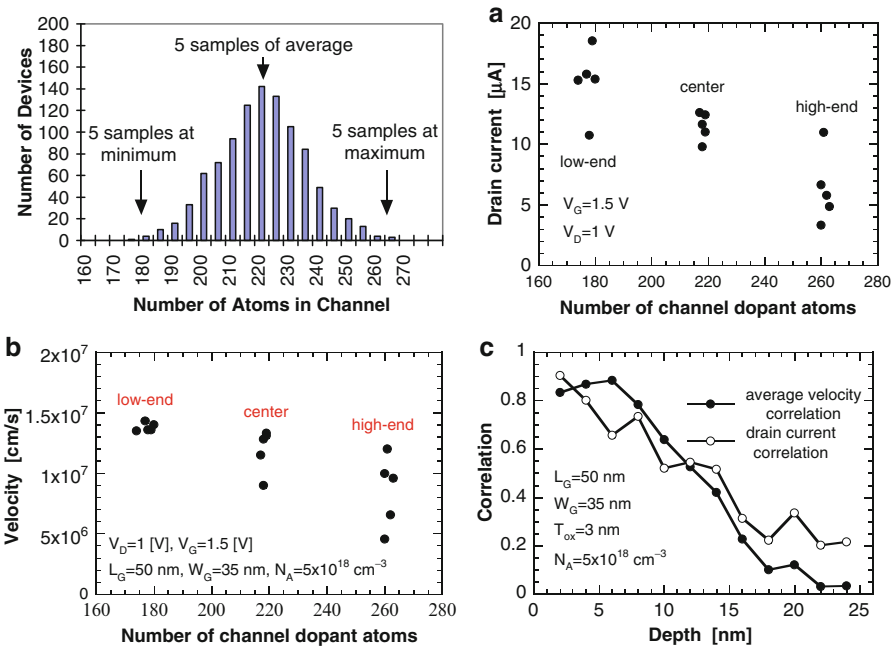


Fig. 2.52 *Top left*: Histogram of the number of dopant atoms in the channel for a population of 1,000 devices. *Top right*: Correlation of the drain current and average electron velocity to the number of dopant atoms within a 10 nm range at various depths beneath the channel. *Bottom left*: Drain current vs. the number of channel dopant atoms. *Bottom right*: Average velocity of channel electrons vs. the number of channel dopant atoms

which the correlation is nearly zero. It appears that only the dopant atoms in the first 6–10 nm from the semiconductor/oxide interface have significant effect on the velocity. This is reinforced by the fact that the correlation nearly goes to zero at a depth of 18 nm, as opposed to the threshold voltage correlation, which remains fairly high at a larger depth. The correlation of the drain current to the number of dopant atoms is also high near the surface, but the drop-off is not as steep as the velocity correlation. Beyond 18 nm depth, the correlation of the drain current is non-zero due to the correlation of the threshold voltage to the number of dopant atoms (see previous discussion).

3.3.2 Threshold Voltage Fluctuations Due to Unintentional Doping in Narrow-Width SOI Device Structures

The SOI device structure that has been simulated in this work to study comprehensively the effects of quantum mechanical size-quantization and discrete/unintentional dopant effects on the performance of nanoscale devices is shown in Fig. 2.53. It consists of a thick (600 nm) silicon substrate, on top of which is grown

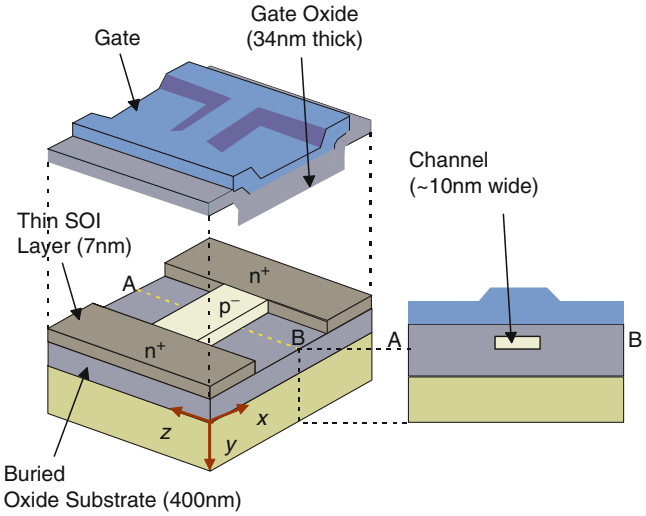


Fig. 2.53 Device structure of ultra-narrow channel FD-SOI device

400 nm of buried oxide. The thickness of the silicon on insulator layer is 7 nm, *with p^- region width of 10 nm* (if not stated otherwise) making it a fully-depleted device under normal operating conditions. The channel length is 50 nm and the doping of the p^- active layer is $10^{16}\,\mathrm{cm}^{-3}$ which corresponds to a nearly undoped channel region. The source/drain length is 15 nm, width being three times the channel width i.e. 30 nm. On top of the SOI layer sits the gate-oxide layer with a thickness of 34 nm. This is rather a thick gate oxide, but it is used to compare the simulation results with the experimental data of Majima et al. [108]. The doping of the source/drain junctions equals $10^{19}\,\mathrm{cm}^{-3}$ (if not stated otherwise), and the gate is assumed to be a metal gate with workfunction equal to the semiconductor affinity. The use of the low source-drain doping is justified by the fact that most of the carriers that are being simulated are residing in the source/drain regions and the reduction of the source/drain doping leads to a smaller ensemble of carriers. It has been found via Silvaco ATLAS Drift-Diffusion simulations of similar device structures that a reduction in the source/drain doping by one order of magnitude leads to approximately 20–30% decrease in the on-state current due to the additional source/drain series resistances.

In a 50 by 10 by 7 nm SOI device structure in Fig. 2.53, with a channel doping of $10^{16}\,\mathrm{cm}^{-3}$, one has merely a single dopant atom in the channel region. Even if the channel is undoped, the unavoidable background doping gives rise to at least one ionized dopant being present at a random location within the channel. Also, if an electron becomes trapped in a defect state at the interface, or in the active silicon body, it will introduce a fixed charge in the channel region. These potential sources of localized single charge will introduce a highly localized barrier to the carrier/current flow. Such a *localized barrier* is shown in Fig. 2.54. The device

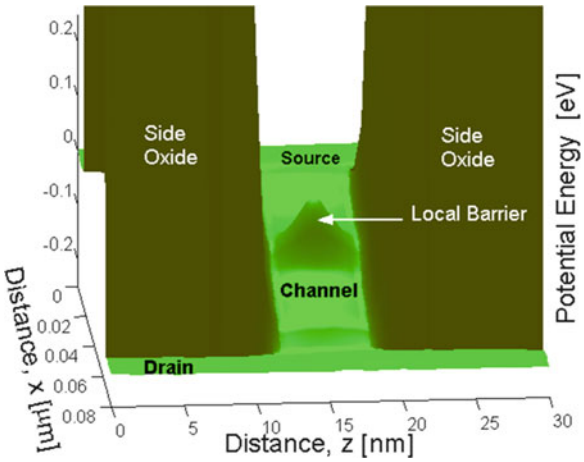


Fig. 2.54 Shape of the conduction band profile when a single impurity is localized in the center of the channel

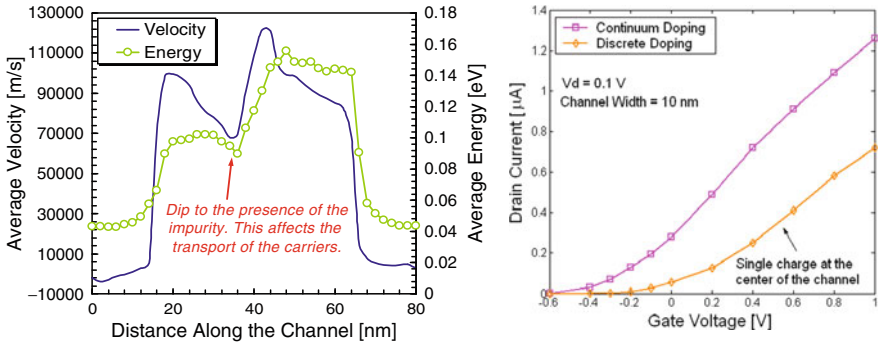


Fig. 2.55 *Left panel*: Velocity and energy plots for $V_G = 1.0$ and $V_D = 0.2$ V when a single impurity is present at the center of the channel. *Right panel*: Device transfer characteristics for the case of a continuum and discrete impurity model with a single charge at the center of the channel

operation is affected by this localized barrier from both electrostatics (effective increase in doping) and dynamics (transport) points of view. The transport is affected through modulation of carrier velocity and energy characteristics as shown in Fig. 2.55 (left panel) where the dip is due to the presence of a single impurity in the center of the channel region. In Fig. 2.55 (right panel), the device transfer characteristics are shown for a device with continuum doping and with an unintentional dopant present in the center of the channel. The channel width is 10 nm. One observes increase in the device threshold voltage V_{th} and degradation of the drain current due to the presence of a single charge.

In Fig. 2.56 shown are the fluctuations in the drain current as a function of the position of a single dopant ion in the channel region of the device. Simulations have

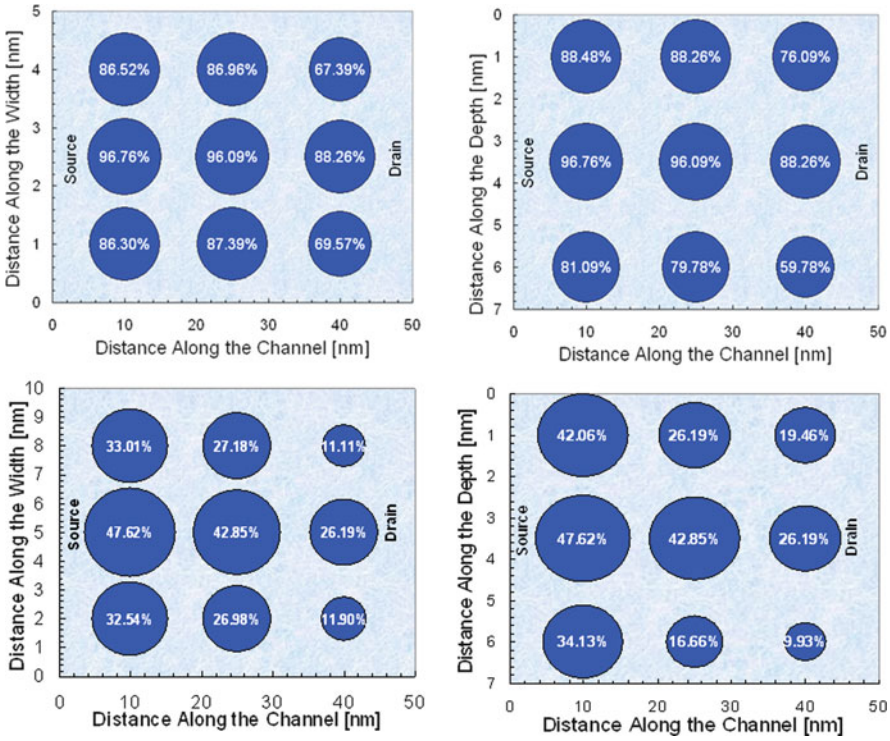


Fig. 2.56 Slicing of the region and corresponding variables in the slices

been performed using $V_G = 1.0$ V and $V_D = 0.1$ V. Results for devices with channel width of both 10 and 5 nm are shown. Due to the size-quantization effect, as a consequence of the charge set-back, results in the majority of current flowing through the middle portion of the channel. Thus a dopant ion trapped in the center region of the channel produces maximum fluctuations in the on-state current. The drain-end is less affected due to two reasons: (a) the presence of a weaker quantization effect therein due to the least vertical field experienced by the electrons and (b) the presence of the largest in-plane (x-component) electric field along the length of channel region which obviously minimizes the effect of the single dopant.

To investigate the impact of screening effect for the impurity positioned along the center of the channel region on the drain current detailed simulations were performed. The results are shown in Fig. 2.56. One can see that the impurity positioned in the very vicinity of the source-end has lower effect than when it is positioned a little away from the source-end. This is attributed to the fact that the very presence of a large number of electrons in the source region try to screen further the impurity and thereby its effect on the drain current.

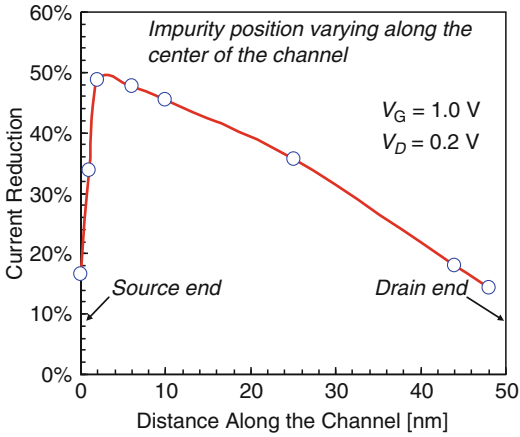


Fig. 2.57 Impact of screening on the drain current

The impurity position dependence of the drain current is shown in Fig. 2.57 (left panel) in the device output characteristics. There are several noteworthy conclusions that can be drawn from these simulations:

- Single impurity at the source-end of the channel affects the drain current the most.
- Impurities at the drain-end of the channel reduce the DIBL (drain-induced-barrier-lowering) in the output characteristics.
- Dopant atoms trapped in the center region of the channel produce the maximum fluctuations than the dopant atoms near the interface.

The observed impurity position dependence of the drain current may be attributed to both the inhomogeneities in the electrostatics and the non-uniform carrier quantization in the channel region. Another potential source arises from the modulation of the transport characteristics, which is reflected in the carrier velocity behavior as shown in the right panel of Fig. 2.58. Here, the velocity profiles for impurities at three different positions are shown. One can see that the impurity near the source end affects (reduces) the electron velocity most, throughout the channel region. Simulations have been performed using $V_G = 1.0\,\mathrm{V}$ and $V_D = 0.2\,\mathrm{V}$.

The results presented in Fig. 2.58 also suggest that there might be fluctuations in the device threshold voltage for devices fabricated on the same chip due to unintentional doping and random positioning of the impurity atoms. This can also be deduced from the scatter of the experimental data from [109]. The simulation results of the transfer characteristics with a single impurity present in different regions in the channel of the device, shown in the left panel of Fig. 2.59 clearly demonstrates the origin of the threshold voltage shifts for devices with 10 and 5 nm channel width. The width dependence of the threshold voltage for the case of a uniform (undoped) and a discrete impurity model is shown in the right panel of Fig. 2.59. This figure suggests that *both size-quantization effects and unintentional doping must be concurrently considered to explain threshold voltage variation in small devises*.

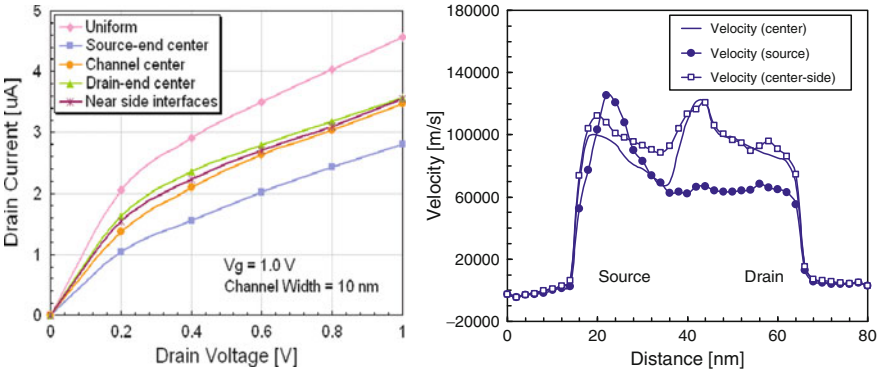


Fig. 2.58 *Left panel*: Variations of the device drain current as a function of the placement of a single impurity at various positions in the channel. We have used $V_G = 1.0\,V$ in these simulations. *Right panel*: Variations of the electron velocity as a function of the placement of a single impurity at various positions in the channel. We have used $V_G = 1.0\,V$ and $V_D = 0.2\,V$ in these simulations

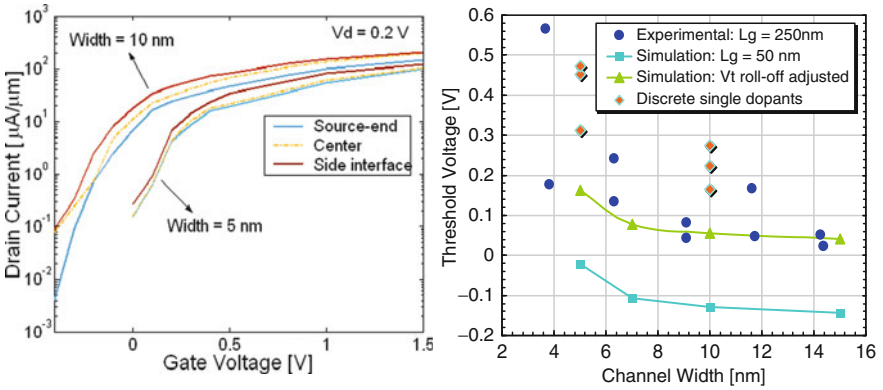


Fig. 2.59 *Left panel*: Transfer characteristics of the device with 10 and 5 nm channel widths and different location of the impurity atoms. We have used $V_G = 1.0\,V$ in these simulations. *Right panel*: Width dependence of the threshold voltage for the case of a uniform and a discrete impurity model. Clearly seen in this figure are two trends: (**a**) Threshold voltage increase with decreasing channel width due to quantum-mechanical size quantization effects, and (**b**) Scatter in the threshold voltage data due to unintentional doping

3.3.3 The Role of Unintentional Doping on FinFET Device Design Parameters

The FinFET device structure that has been simulated in this work is shown in Fig. 2.60 [109]. It consists of a thick (100 nm) buried oxide on top of which source/drain regions and a vertical fin are formed. The channel length is 40 nm with a gate length of 20 nm and a fin extension length of 10 nm on each side of the gate. The fin height and width are 30 and 10 nm, respectively. The source/drain length

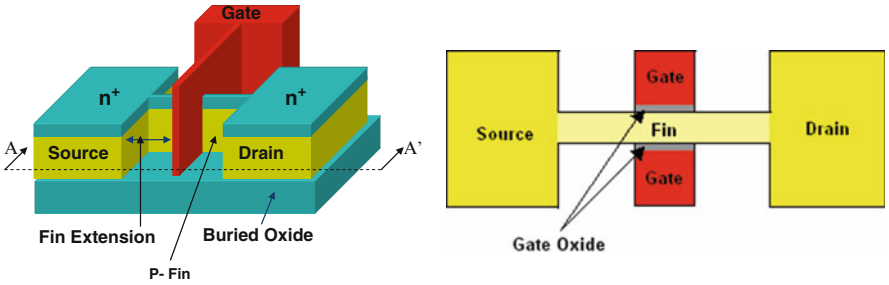


Fig. 2.60 *Left panel*: 3D schematic view of FinFET. *Right panel*: Top view of the FinFET shown in top panel along the cross section A-A'

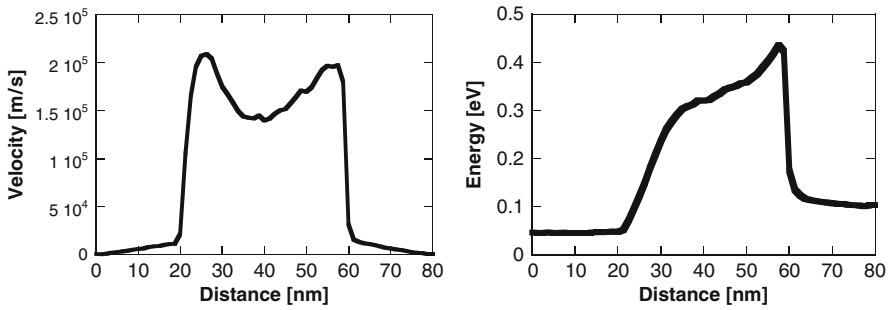


Fig. 2.61 *Left panel*: Average velocity (x-component) profile of carriers along the channel. *Right panel*: Average energy of carriers along the length of the device. $V_G = V_D = 0.8\,V$ and $S/G = D/G = 10\,nm$

is 20 nm, the width being three times the channel width, i.e. 30 nm. The doping of the source/drain junctions equals $2 \times 10^{19}\,cm^{-3}$. The fin is assumed intrinsic. The gate is assumed to be n^+ polysilicon with work function equal to the semiconductor affinity. Gate oxide of 2.5 nm has been used for both side and top gates. To simulate this device structure, a convenient meshing scheme has been adopted. Meshing is uniform along the x (channel length) and z (width) directions and is non-uniform along the y (depth) direction, with the exception of the semiconductor region, where uniformity in meshing has been kept in order to facilitate the Monte Carlo transport simulations.

Significant velocity overshoot is observed in small geometry devices due to the presence of very high electric fields. Figure 2.61 (left panel) depicts the average velocity profile along the channel length of a FinFET device. Equal amount of velocity overshoot is observed near the source and the drain end of the channel when fin extension length on each side of the gate is equal. Note that the magnitude of the velocity overshoot also depends on the fin extension length on each side of the gate and this observation is discussed later in the text. Figure 2.61 (right panel) depicts the average energy profile along the device channel length. Near the source end the average carrier energy equals the thermal energy. Along the channel the average

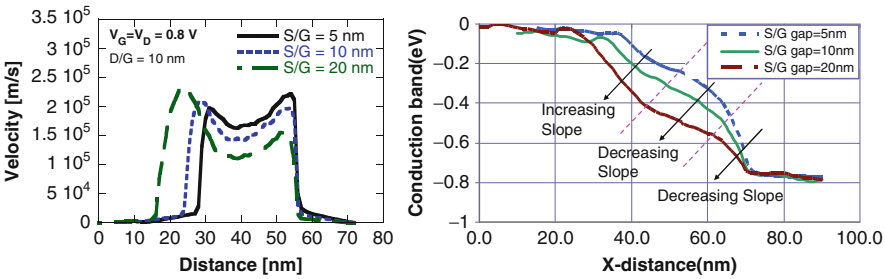


Fig. 2.62 *Left panel*: Average velocity (x-component) profile of carriers along the channel as a function of S/G gap. The applied bias equals $V_{\mathrm{G}} = V_{\mathrm{D}} = 0.8\,\mathrm{V}$. *Right panel*: Conduction band profile along x-direction

energy increases progressively reaching its peak value near the drain end. Note that carriers are not thermalized near the drain end of the channel due to the omission of the short-range electron–electron and electron–ion interactions in these simulations. Fin extension of 10 nm has been used on each side of the gate. The applied bias equals $V_D = V_G = 0.8\,\mathrm{V}$.

The amount of velocity overshoot the carriers experience within the FinFET devices shown previously heavily depends on the fin extension length on each side of the gate. Keeping D/G gap fixed, gradual increase in S/G gap causes the source end to experience more overshoot and the drain side overshoot to gradually diminish as shown in Fig. 2.62 (left panel). This is due to the fact that with an increase in extension length, source and drain lateral fields along the channel redistribute which changes the velocity profiles which can be seen from the 1-D conduction band profile along the x-direction as shown in Fig. 2.62 (right panel). Near the drain end and in the channel the slope of conduction band decreases with increase in S/G gap, resulting in lower electric field. Also note that near the source end the slope of conduction band increases giving higher electric field at that region. D/G gap is fixed at 10 nm and $V_D = V_G = 0.8\,\mathrm{V}$ is used in the simulation. The same phenomena happen for varying the D/G gap while keeping S/G gap constant at 10 nm.

From the transfer characteristics of the device as shown in Fig. 2.63 (left panel), it is evident that the threshold voltage is negative and is around $-0.1\,\mathrm{V}$. Negative threshold voltage results due to the use of n^{+}-polysilicon as a gate electrode. The metal work function equal to the electron affinity of Si is assumed in the simulation. Polysilicon gates also suffer from depletion and high gate resistance. A nominal threshold voltage of 0.2–0.4 V for n-channel FinFET can be achieved using metal gates with work function close to the mid band-gap of silicon ($\sim 4.6\,\mathrm{e}V$). Achieving symmetric threshold voltages for both n-channel and p-channel FinFETs requires metals with different work functions [110]. The output characteristics of the device from Fig. 2.60 are presented in Fig. 2.63 (right panel). Equal fin extension of 10 nm is assumed on both sides of the gate. Gate voltage $V_G = 0.4\,\mathrm{V}$ is used. The inclusion of the electron–electron and electron–ion interaction results in lower drain current.

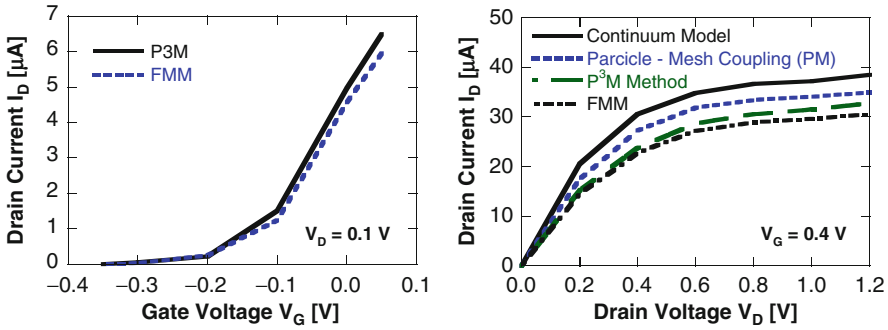


Fig. 2.63 *Left panel*: Transfer characteristics. *Right panel*: Output characteristics

Table 2.3 P^3M vs. FMM speed-up

Approach	CPU time per iteration (s)
P^3M	∼24
FMM	<1

Also the Fast Multipole method (FMM) gives output characteristic which is in good agreement with that using the P^3M approach.

It is important to note that the CPU time requirement when using the FMM is much smaller compared to the traditional P^3M approach. Table 2.3, gives a comparison of the CPU time requirements for simulating FinFET device with a 3D mesh of $64 \times 24 \times 24$ node points. The number of particles simulated is around 1,500. The speedup due to using FMM depends on the number of particles, mesh size and computational resources. As the number of particles increases, FMM becomes slower but still much faster when compared to the P3M approach. Also for very small number of particles, it is better to calculate e–e and e–ion interaction directly than using FMM [111]. Correction for image charges is incorporated in our simulator to get the precise results.

FinFET devices use undoped or lightly doped fin. In a 40 by 10 by 30 nm channel region, with a channel doping of $10^{16}\,\mathrm{cm}^{-3}$, one has merely 0.12 dopant atoms in the channel region. Even if the channel is undoped, the unavoidable background doping gives rise to at least one ionized dopant being present at a random location within the channel. Also, if an electron becomes trapped in a defect state at the interface or in the silicon body, it will introduce a fixed charge in the channel region. These potential sources of localized single charge will introduce a localized barrier to current flow. The position of a single dopant at the center of the channel along with the localized barrier it creates is shown in Fig. 2.64 (left and right panel). The device operation is affected by this localized barrier from both electrostatics (effective increase in doping) and dynamics (transport) points of view. The effective increase in doping in the channel region results in increase in the threshold voltage and consequently, the drain current reduces. The transport is affected through modulation of the carrier velocity and energy characteristics.

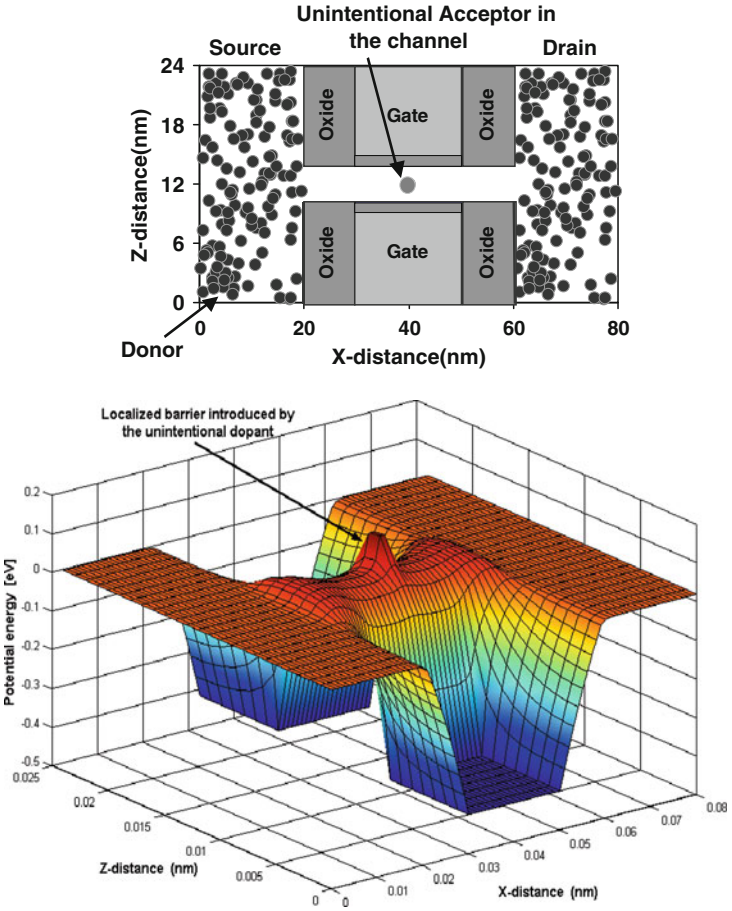


Fig. 2.64 *Left panel*: Top view of the FinFET device showing dopant position at the center region of the channel. *Right panel*: Potential profile showing the localized barrier introduced by the unintentional dopant

Due to the presence of multiple channels in the FinFET device, the effect of unintentional doping is not that much pronounced. The reduction in drain current heavily depends on the fin width. With decrease in fin width, the localized barrier has more pronounced effect on carrier motion through the channel, and the reduction in drain current is significant. This trend is schematically shown on the left panel of Fig. 2.65 . Fin extension length of 10 nm is used on each side of the gate. $V_D = 0.1\,\mathrm{V}$, $V_G = 0.4\,\mathrm{V}$ is used in the simulation. The unintentional dopant is placed near the source end close to the top interface. Fin extension length on each side also influences the reduction in drain current due to unintentional dopant as it is shown

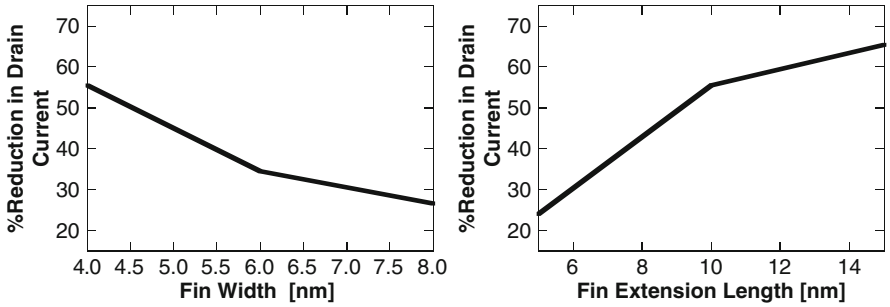


Fig. 2.65 *Left panel*: Reduction in drain current due to unintentional dopant as a function of fin width. $V_G = 0.4\,V$, $V_D = 0.1\,V$. *Right panel*: Reduction in drain current due to unintentional dopant as a function of fin extension length. $V_G = 0.4\,V$, $V_D = 0.1\,V$

in the right panel of Fig. 2.65. Longer fin extension results in more reduction in drain current than that due to smaller fin extension for any dopant position. With longer fin extension, lateral field from source and drain has less influence on the barrier produced by the unintentional dopant thereby, reducing the drain current more when compared to the case with smaller fin extension. Fin extension length can therefore, be optimized for suppressing unintentional doping effects while keeping the drive current within required range. $V_G = 0.4\,V$ and $V_D = 0.1\,V$ is used. The dopant atom is placed near the source end close to the top interface. Fin width of 4 nm is used. As noted in earlier device structures, the reduction in drain current due to unintentional dopant significantly depends on the position of the dopant atom in the channel. It is found that dopant placed near the source end has greater effect on the drain current. Near the drain end, the effect is less pronounced. Since in FinFET devices channels are formed symmetrically in vertical plane on each side of the fin, placing the unintentional dopant near the center along the width will reduce drain current more than that caused by dopant for any other position.

The effect of unintentional doping on device operation is relatively strong near sub threshold regime/weak inversion when few carriers are present in the channel. Thus the presence of unintentional dopant in the channel is expected to affect the switching behavior of the device. Increasing either the gate voltage or the drain bias will reduce the effect. As the gate voltage is increased, the number of carriers in the channel region increases and screens the localized potential produced by the unintentional dopant as shown in the left panel of Fig. 2.66. Drain bias of 0.1 V is applied in the simulation. Unintentional dopant is placed at the center of the channel near the top interface. Similarly with increase in drain voltage carriers are accelerated more along the channel and thus, can easily overcome the localized barrier. Therefore the reduction in drain current gradually decreases with increasing drain bias as shown in the right panel of Fig. 2.66. Gate bias of 0.4 V is applied in the simulation. Dopant is placed near the source end of the fin close to the top interface.

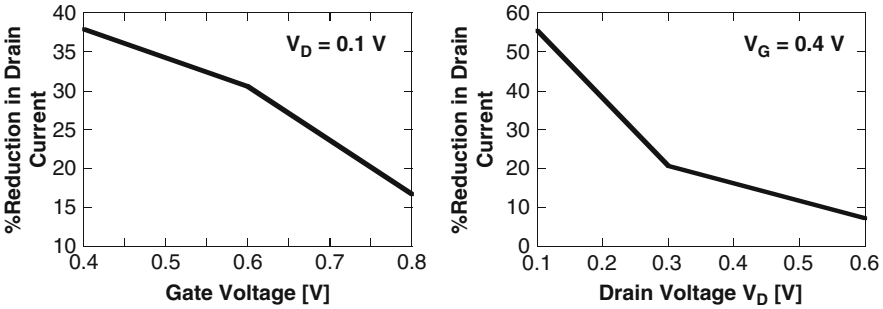


Fig. 2.66 *Left panel*: Screening behavior of the carriers on reduction of drain current due to unintentional dopant. *Right panel*: Reduction in drain current due to unintentional dopant as a function of drain voltage

4 Conclusions

A recently proposed *effective potential* approach has been utilized to successfully simulate two-dimensional space-quantization effects in a model of a narrow-channel SOI device structure. The incorporation of the *effective potential* approach into a full 3D Monte Carlo particle-based simulator allows one to investigate the device transfer and output characteristics with proper treatment of the size-quantization effects, velocity overshoot and carrier heating on an equal footing. The *effective potential* provides a set-back of the charge from the interface proper and quantization energy within the channel. Both of these effects lead to an increase in the threshold voltage. A threshold voltage increase of about 180 mV has been observed when the effective potential is included in the SOI device with 10 nm channel width. Also, observed is a pronounced channel width dependency of the threshold voltage which is termed as the *quantum mechanical narrow channel effect*. The width dependence of the threshold voltage is in close agreement with the experimental results. The increase in the threshold voltage is found to give rise to a significant on-state current reduction (20–30%), which depends upon the gate bias. Larger degradation is observed for larger gate voltages. The energy characteristics along the channel do not change with the inclusion of quantum mechanical size-quantization effects. The average drift velocity shows a small decrease due to the smearing of the potential.

A novel effective potential approach has been proposed and tested in the simulations of quantization effects in 25 nm nano-MOSFET device. The approach is parameter free as the size of the electron depends upon its energy. We have justified the correctness of the approach with simulations of the gate voltage dependence of the sheet electron density. The excellent agreement between the simulations and SCHRED results suggests that one is able to correctly predict the effective oxide thickness increase due to quantum-mechanical size-quantization effects that leads to a reduction of the sheet electron density. The nano-MOSFET simulation results also confirm this charge displacement effect near the source end of the channel where quantization effects play significant role. Due to the larger smearing of the

potential for high energy electrons, we see a decrease in the carrier velocity when quantization effects are included in the model. This leads to a smaller drain current in both the device transfer and output characteristics. The charge displacement from the interface, and the effective increase of the oxide thickness, gives rise to a threshold voltage shift of $\sim$220 mV which is consistent with earlier observations. The shift in the threshold voltage leads in turn, to a drain current degradation of about 30%. Hence, the observations presented here that utilize the new effective potential approach, confirm that quantum-mechanical space-quantization effects must be included in the theoretical model to correctly predict the device behavior. In some cases, this can be achieved with the incorporation of the barrier field that is precomputed in the initial stages of the simulation and does not require additional CPU time during the simulation sequence. We believe that this new effective potential approach is more reliable in simulation of quantization effects in nano-scale devices with barriers that have different size and shape.

To treat the short-range Coulomb (electron–ion and electron–electron) interactions properly, *three* different but consistent real-space *molecular dynamics* (MD) schemes have been implemented in the simulator: the particle-particle-particle-mesh (P^3M) method, the corrected Coulomb approach and the Fast Multipole Method (FMM). *It is believed that the FMM algorithm has been used for the first time in the simulations of semiconductor devices*. The correctness of the approaches is verified via the simulations of the doping dependence of the low-field electron mobility in a 3D resistor and through its comparison with available experimental data. These approaches are then applied in the investigations of the role of unintentional doping on the operation of narrow-width SOI devices. We find significant correlation between the location of the impurity atom and the magnitude of the drain current. Namely, impurities near the source end of the channel have maximum influence on the drain current. This observation suggests that one has to take into account transistor mismatches due to unintentional doping when performing circuit designs. We have also investigated in depth the fluctuations in the threshold voltage due to discrete distribution of the impurity atoms in narrow width SOI devices with 10 and 5 nm channel width. The simulated data for the threshold voltages are in perfect agreement with the experimental values and they explain the fluctuations in the experimentally derived threshold voltage data.

Another device structure that has been investigated regarding the influence of the discrete impurities is the FinFET. Among different double gate structures FinFET attracts the researchers due to its inherent immunity to short channel effects and ease of fabrication using the existing planar fabrication process flow. Single fin FinFET can easily be extended to multiple fin structure for higher drive current. Again, in this structure as well, we find significant correlation between the magnitude of the drain current and the position of the discrete dopant for the case when screening effects do not play considerable role.

Acknowledgements We would like to thank the financial support from the National Science Foundation under Contract Number ECCS 0901251: Modeling Heating Effects in Low-Power Multi-Gate SOI Devices and High-Power GaN HEMTs. Program Director: Paul Werbos.

References

1. W. Hansch, Th. Vogelsang, R. Kirchner and M. Orlowski., "Carrier Transport Near the Si/SiO2 Interface of a MOSFET", *Solid State Elec.*, vol 32, no. 10, pp. 839–849, Oct. 1989.
2. M.J. Van Dort, PH. Woerlee and AJ. Walker, "A Simple Model for Quantization Effects in Heavily-Doped Silicon MOSFETs at Inversion Conditions.", *Solid State Elec.*, vol. 37, no. 3, pp. 411–415, Mar. 1994.
3. F. F. Fang and W. E. Howard, "Negative Field-Effect Mobility on (100) Si Surfaces", *Phys. Rev. Lett.*, vol. 16, no. 18, pp. 797–799, May. 1966.
4. B. Winstead and U. Ravaioli, "Simulation of Schottky barrier MOSFET's with a coupled quantum injection/Monte Carlo technique," *IEEE Trans. Electron Devices*, vol. 47, no. 6, pp. 1241–1246, Jun. 2000.
5. R. W. Keyes, "The effect of randomness in the distribution of impurity atoms on FET thresholds," *Appl. Phys.*, vol. 8, no. 3, pp. 251–259, Jun. 1975.
6. T. Mizuno, J. Okamura, and A. Toriumi, "Experimental study of threshold voltage fuctuation due to statistical variation of channel dopant number in MOSFET's," *IEEE Trans. Electron Devices*, vol. 41, pp. 2216–2221, Nov. 1994.
7. H. S. Wong and Y. Taur, "Three dimensional 'atomistic' simulation of discrete random dopant distribution effects in sub-0.1 mm MOSFET's, in *IEDM Tech. Dig.*, pp. 705–708, Dec. 1993.
8. W. J. Gross, D. Vasileska, and D. K. Ferry, "3-D Simulations of ultrasmall MOSFET's with real-space treatment of the electron–electron and electron–ion interactions," *VLSI Design*, vol. 10, pp. 437–452, no. 4, 2000.
9. A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub 0.1 micron MOSFETs: A 3D 'atomistic' simulation," *IEEE Trans. Electron Devices*, vol. 45, no. 12, pp. 2505–2513, Dec. 1988.
10. William J. Gross, *Ph. D. Dissertation*, Arizona State University, Dec. 2000.
11. N. Sano, K. Matsuzawa, M. Mukai, and N. Nakayama, "Role of longrange and short-range Coulomb potentials in threshold characteristics under discrete dopants in sub-0.1um Si-MOSFETs," *IEDM Tech., Dig.*, pp. 275–283, Dec. 2000.
12. D. K. Ferry, A. M. Kriman, M. J. Kann, and R. P. Joshi, "Molecular dynamics extensions of Monte Carlo simulation in semiconductor device modeling", *Comp. Phys. Comm.*, vol. 67, no. 1, pp. 119–134, Aug. 1991.
13. L. R. Logan and J. L. Egley, "Dielectric response in p-type silicon: Screening and band-gap narrowing", *Phys. Rev. B*, vol. 47, no. 19, pp. 12532–12539, May. 1993.
14. C. Jacoboni and P. Lugli, *"The Monte Carlo Method for Semiconductor Device Simulation."*, Vienna, Austria: Springer-Verlag, 1989.
15. R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous and A. R. leBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions", *IEEE J. Solid-State Circuits*, vol. 9, pp. 256, 1974.
16. J. R. Brews, W. Fichtner, E. H. Nicollian and S. M. Sze, "Generalized guide for MOSFET miniaturization", *IEEE Electron Dev. Lett.*, vol 1, no. 2, pp. 2, Jan. 1980.
17. G. Bacarani and M. R. Worderman, "Transconductance degradation in thin-Oxide MOSFET's", *Electron Devices Meeting*, pp. 278–281, (1982).
18. M.-S. Liang, J. Y. Choi, P.-K. Ko and C. Hu, "Inversion-Layer Capacitance and Mobility of Very Thin Gate-Oxide MOSFET's", *IEEE Trans. Electron Devices*, vol. 33, no. 3, pp. 409–413, Mar. 1986.
19. A. Hartstein and N. F. Albert, "Determination of the inversion-layer thickness from capacitance measurements of metal-oxide-semiconductor field-effect transistors with ultrathin oxide layers", *Phys. Rev. B*, vol. 38, no. 2, pp. 1235–1240, Jul. 1988.
20. M. J. van Dort, P. H. Woerlee, A. J. Walker, C. A. H. Juffermans and H. Lifka, "Influence of high substrate doping levels on the threshold voltage and the mobility of deep-submicrometer MOSFETs", *IEEE Trans. Electron Dev.*, vol. 39, no. 4, pp. 932–938, 1992.
21. M. J. van Dort, P. H. Woerlee and A. J. Walker, "A simple model for quantisation effects in heavily-doped silicon MOSFETs at inversion conditions", *Solid-State Electronics* **37**, 411 (1994).

22. D. Vasileska, and D.K. Ferry, "The influence of space quantization effects on the threshold voltage, inversion layer and total gate capacitance in scaled Si-MOSFETs," *Technical Proceedings of the First International Conference on Modeling and Simulation of Microsystems, Semiconductors, Sensors and Actuators, Santa Clara, California*, pp. 408–413, Apr. 1998.
23. S. Takagi and A. Toriumi, "Quantitative understanding of inversion-layer capacitance in Si MOSFET's", *IEEE Trans. Electron Devices*, vol. 42, no. 12, pp. 2125–2130, Dec. 1995.
24. S. A. Hareland, S. Krishnamurthy, S. Jallepali, C. F. Yeap, K. Hasnat, A. F. Tasch Jr. and C. M. Maziar, "A computationally efficient model for inversion layer quantization effects in deep submicron N-channel MOSFETs", *IEEE Trans. Electron Devices*, vol. 43, no. 1, pp. 90–96, Jan. 1996.
25. D. Vasileska, D. K. Schroder and D. K. Ferry, "Scaled silicon MOSFETs: degradation of the total gate capacitance", *IEEE Trans. Electron Devices*, vol. 44, no. 4, pp. 584–587, 1997.
26. K. S. Krisch, J. D. Bude and L. Manchanda, "Gate capacitance attenuation in MOS devices with thin gate dielectrics", *IEEE Electron Dev. Lett*, vol. 17, no. 11, pp. 521–524, Nov. 1996.
27. L. de Broglie, *C. R. Acad. Sci. Paris*, vol. 183, 447, 1926.
28. L. de Broglie, *C. R. Acad. Sci. Paris*, vol. 184, 273, 1927.
29. E. Madelung, "Quantum theory in hydrodynamical form", *Z. Phys.*, 40, 322, 1926.
30. D. Bohm, "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I", *Phys. Rev.*, 85, no. 2, 166–179, Jan. 1952.
31. D. Bohm, "A suggested interpretation of the quantum theory in terms of hidden variables. II", *Phys. Rev.*, Vol. 85, 180 (1952).
32. C. Dewdney and B. J. Hiley, "A Quantum Potential Description of One-Dimensional Time-Dependant Scattering From Square Barriers and Square Wells", *Found. Phys.*, vol. 12, no. 1, pp. 27–48, Jan. 1982.
33. G. J. Iafrate, H. L. Grubin, and D.K Ferry, "Utilization of Quantum Distribution Functions for Ultra-Submicron Device Transport", *Journal de Physique.*, vol. 42 (Colloq. 7), 10, 307–312, Oct. 1981.
34. E. Wigner, "On the Quantum Correction For Thermodynamic Equilibrium", *Phys. Rev.*, vol. 40, no. 5, pp. 749–759, Jun. 1932.
35. D. K. Ferry and J.-R. Zhou, "Form of the quantum potential for use in hydrodynamic equations for semiconductor device modeling", *Phys. Rev. B.*, vol. 48, no. 11, pp. 7944–7950, Sep. 1993.
36. P. Feynman and H. Kleinert, "Effective classical partition functions", *Phys. Rev. A*, 34, no. 6, pp. 5080–5084, Dec. 1986.
37. C. L. Gardner and C. Ringhofer, "Smooth quantum potential for the hydrodynamic model", *Phys. Rev. E*, vol. 53, no. 1, pp. 157–166, Jan. 1996.
38. C. Ringhofer and C. L. Gardner, "Smooth quantum hydrodynamic model simulation of the resonant tunneling diode", *VLSI Design*, vol. 8, 1–4, 143–146, 1998.
39. D. Vasileska and S. S. Ahmed, "Narrow-Width SOI Devices: The Role of Quantum–Mechanical Size Quantization Effect and Unintentional Doping on the Device Operation", *IEEE Trans. Electron Devices*, vol. 52, no. 2, pp. 227–236, Feb. 2005.
40. D. K. Ferry, "The onset of quantization in ultra-submicron semiconductor devices", *Superlattices and Microstructures*, vol. 27, no. 2–3, pp. 61–66, Jan. 2000.
41. C. Ringhofer, S. Ahmed and D. Vasileska, "An effective potential approach to modeling 25 nm MOSFET devices", *Journal of Computational Electronics*, vol. 2, pp. 113–117, 2003.
42. C. Ringhofer, C. Gardner and D. Vasileska, "Effective potentials and quantum fluid models: a thermodynamic approach", *Inter. J. on High Speed Electronics and Systems*, vol. 13, no. 3, pp. 771–804, Jan. 2003.
43. Shaikh Shahid Ahmed, "Quantum and Coulomb Effects in Nanoscale Devices", Ph. D. Dissertation, *Arizona State University*, Dec. 2004.
44. R. Akis, S. Milicic, D. K. Ferry, D. Vasileska, "An Effective Potential Method for Including Quantum Effects Into the Simulation of Ultra-Short and Ultra-Narrow Channel MOSFETs", *Proceedings of the 4th International Conference on Modeling and Simulation of Microsystems, Hilton Head Island, SC*, pp. 550–3, Mar. 2001.

45. C. Ringhofer, S. S. Ahmed and D. Vasileska, "Effective potential approach to modeling of 25 nm MOSFET devices", *Superlattices and Microstructures*, vol. 34, no. 3–6, pp. 311–317, 2003.
46. http://www.intel.com
47. Y. Omura, S. Horiguchi, M. Tabe, and K. Kishi, "Quantum-mechanical effects on the threshold voltage of ultrathin-SOI nMOSFETs", *IEEE Elec. Device Lett.*, vol. 14, no. 12, pp. 569–571, Dec. 1993.
48. S. M. Ramey and D. K. Ferry, "Implementation of surface roughness scattering in Monte Carlo modeling of thin SOI MOSFETs using the effective potential", *IEEE Transactions on Nanotechnology*, vol. 2, no. 2, pp. 110–114, Jun. 2003.
49. S. Hasan, J. Wang, and M. Lundstrom, "Device design and manufacturing issues for 10 nm-scale MOSFETs: a computational study", *Solid–State Elect*, vol. 48, no. 6, pp. 867–875, 2004.
50. S. Datta, *"Electronic Transport in Mesoscopic Systems"*, Cambridge Studies in Semiconductor Physics Series, ISBN 0-521-59943-1, paperback, 1998.
51. D. Vasileska, S. M. Goodnick and Gerhard Klimeck, *Computational Electronics: Semiclassical and Quantum Transport Modeling*, CRC Press, June 2010.
52. P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", *Phys. Rev.*, vol. 136, no. 3b, B864-B871, Nov. 1964.
53. Kohn, and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", *Phys. Rev.*, vol. 140, no. 4a, pp. A1133–A1138, Nov. 1965.
54. L. Hedin and B. I. Lundqvist, "Explicit local exchange-correlation potentials", *J. Phys. C*, vol. 4, no. 14, pp. 2064–2082, Mar. 1971.
55. C. Hu, S. Banerjee, k. Sadra, B.G. Streetman and R. Sivan, "Quantization Effects in Inversion Layers of PMOSFET's on Si (100) Substrates", *IEEE Electron Dev. Lett.*, vol. 17, no. 6, pp. 276–278, Jun. 1996
56. S. Takagi, M. Takayanagi, and A. Toriumi, "Characterization of Inversion-Layer Capacitance of Holes in Si MOSFET's", *IEEE Trans. Electron Devices*, vol. 46, no. 7, pp. 1446–1450, Jul. 1999.
57. D. Vasileska, D. K. Schroder and D.K. Ferry, "Scaled silicon MOSFET's: Part II-Degradation of the total gate capacitance", *IEEE Trans. Electron Devices*, vol. 44, no. 4, pp. 584–587, Apr. 1997.
58. D. Vasileska, and D.K. Ferry, "The influence of space quantization effects on the threshold voltage, inversion layer and total gate capacitance in scaled Si-MOSFETs", *Technical Proceedings of the First International Conference on Modeling and Simulation of Microsystems, Semiconductors, Sensors and Actuators, Santa Clara, California*, vol. 10, no. 2, pp. 408–413, Apr. 1998.
59. J. Fossum, Z. Ren, K. Kim and M. Lundstrom "Extraordinarily High Drive Currents in Asymmetrical Double-Gate MOSFETs", *Superlattices and Microstructures*, vol. 28, no. 5–6, pp. 525–530, Jun. 2000.
60. J. P. Colinge, X. Baie, V. Bayot, and E. Grivei, "A silicon-on-insulator quantum wire," *Solid-State Electron.*, vol. 39, no. 1, pp. 49–51, Jan. 1996.
61. X. Huang, W. C. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y. K. Choi, K. Asano, V. Subramanian, T. J. King, J. Bokor, and C. Hu, "Sub 50-nm FinFET: PMOS," in *IEDM Tech. Dig.*, pp. 67–70, Dec. 1999.
62. Z. Jiao and C. A. T. Salama, "A fully depleted _-channel SOI nMOSFET," *Electrochem. Soc. Proc.*, vol. 3, pp. 403–408, 2001.
63. J. P. Coolinge, M. H. Gao, A. Romano, H. Maes, and C. Claeys, "Silicon- on-insulator "gate-all-around" MOS device," *SOI Conf. Dig.*, pp. 137–138, 1990.
64. D. Hisamoto, T. Kaga, Y. Kawamoto, and E. Takeda, "A fully depleted lean-channel transistor (DELTA)—A novel vertical ultra-thin SOI MOSFET," in *IEDM Tech. Dig.*, pp. 833–836, Dec. 1989.
65. C. P. Auth and J. D. Plummer, "A simple model for threshold voltage of surrounding-gate MOSFETs,", *IEEE Trans. Electron Devices*, vol. 45, no.11, pp. 2381–2383, Nov. 1998.

66. T. Sekigawa and Y. Hayashi, "Calculated threshold voltage characteristics of an XMOS transistor having an additional bottom gate," *Solid- State Electron.*, vol. 27, no. 8–9, pp. 827–828, Jan. 1984.
67. A. Rahman, M. S. Lundstrom, and A. W. Ghosh, "Generalized effective-mass approach for n-type metal-oxide-semiconductor field-effect transistors on arbitrarily oriented wafers", *Journal of applied physics*, vol. 97, no. 5, pp. 053702–053714, Feb. 2005.
68. A. Rahman, "Exploring new channel materials for nanoscale CMOS devices: A simulation approach", Ph.D. Dissertation, Purdue University.
69. I. H. Tan, G.L. Snider, L. D. Chang and E. L. Hu, "A self-consistent Solution of Schrödinger–Poisson Equations using a Non-uniform Mesh," *J. Appl. Phys.*, vol. 68, pp. 4071–4076, Oct. 1990.
70. T. Yang, Y. Liu, P.D. Ye, Y. Xuan, H. Pal, M. S. Lundstrom, "Inversion Capacitance-Voltage Studies on GaAs Metal-Oxide-Semiconductor Structure using Transparent Conducting Oxide as Metal Gate", *Applied Physics Letters*, vol. 92, pp. 252105–252108, Jun. 2008.
71. F. Gilibert, D. Rideau, F. Payet, F. Boeuf, E. Batail, M. Minondo, R. Bouchakour, T. Skotnicki, H. Jaouen, "Strained Si/SiGe MOSFET capacitance modeling based on band structure analysis", *Proceedings of the 35th European Solid State Device Research Conference (ESSDERC'2005), Grenoble*, no. 12–16, pp. 281–284, Sep. 2005.
72. L. Rayleigh, "On the propagation of waves through a stratified medium, with special reference to the question of reflection", *Proc. Roy. Soc. A*, vol. 86, no. 586, pp. 207, 1912.
73. ET. Jaynes. Probability Theory: The Logic of Science, Cambridge University Press, (2003).
74. Usuki T., Saito M., Takatsu M., Kiehl R.A., Yokoyama N.:*Numerical analysis of electron wave detection by a wedge shaped point contact*. Phys. Rev. B 520, 7615–7625 (1994).
75. Datta S.: *Nanoscale device modeling: the Green's function method*. Superlattices and Microstuctures 28, 253–278 (2000).
76. D. K. Ferry, *Quantum Mechanics for Electrical Engineers*, IOP Press (2000).
77. C. B. Duke, in Solid State Physics, edited by F. Seitz, D. Turnbull, and H. Ehrenreich ∼ Academic, New York, 1969!
78. W. W. Lui and M. Fukuma "Exact solution of the Schrödinger equation across an arbitrary one-dimensional piecewise-linear potential barrier" *J. Appl. Phys*. 60, 1555–1559 (1986).
79. T. Mizuno, J. Okamura and A. Toriumi, "Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in MOSFET's", *IEEE Trans. Electron Devices*, vol. 41, no. 11, pp. 2216–2221, Nov. 1994.
80. T. Mizuno, "Influence of Statistical Spatial-Nonuniformity of Dopant Atoms on Threshold Voltage in a System of Many MOSFETs", *Jpn. J. Appl. Phys*, vol. 35, pp. 842–848, Jan. 1996.
81. J. T. Horstmann, U. Hilleringmann and K. F. Goser, "Matching analysis of deposition defined 50-nm MOSFET's", *IEEE Trans. Electron Devices*, vol. 45, no. 1, pp. 299–306, Jan. 1998.
82. P. A. Stolk, F. P. Widdershoven and D. B. M. Klaassen, "Modeling statistical dopant fluctuations in MOS transistors", *IEEE Trans. Electron Devices*, vol. 45, pp. 1960–1971, Sep. 1998.
83. K. Nishinohara, N. Shigyo and T. Wada, "Effects of microscopic fluctuations in dopant distributions on MOSFET threshold voltage", *IEEE Trans. Electron Devices*, vol. 39, no. 3, pp. 634–639, Mar. 1992.
84. J.-R. Zhou and D. K. Ferry, "3D simulation of deep-submicron devices. How impurity atoms affect conductance", *IEEE Comput. Science and Eng.*, vol. 2, no. 2, pp. 30–36, May. 1995.
85. D. Vasileska, W. J. Gross, V. Kafedziski and D. K. Ferry, "Continuity Equations for Scaled Si MOSFETs", *VLSI Design*, vol. 8, no. 1–4, pp. 301, 1998.
86. D. Vasileska, W. J. Gross and D. K. Ferry, "Modeling of deep-submicrometer MOSFETs: random impurity effects, threshold voltage shifts and gate capacitance attenuation", *Extended Abstracts IWCE-6, Osaka, IEEE Cat*. No. 98EX116, pp. 259–262, 1998.
87. X. Tang, V. K. De and J. D. Meindl, "Intrinsic MOSFET parameter fluctuations due to random dopant placement", *IEEE Trans. on VLSI Systems*, vol. 5, no. 4, pp. 369–376, Dec. 1997.

88. P. Lugli and D. K. Ferry, "Degeneracy in the ensemble Monte Carlo method for high-field transport in semiconductors", *IEEE Trans. Electron Dev.*, vol. 32, no. 11, pp. 2431–2437, Nov. 1985.
89. A. M. Kriman, M. J. Kann, D. K. Ferry and R. Joshi, "Role of the exchange interaction in the short-time relaxation of a high-density electron plasma", *Phys. Rev. Lett.*, vol. 65, no. 13, pp. 1619–1622, Sep. 1990.
90. W. J. Gross, D. Vasileska, and D. K. Ferry, "3 D simulations of ultra-small MOSFETs with real-space treatment of the electron-electron and electron-ion interactions", *VLSI Design*, vol. 10, no. 4, pp. 437-, 2000.
91. D. Vasileska, W. J. Gross, and D. K. Ferry, "Monte Carlo particle-based simulations of deep-submicron n-MOSFETs with real-space treatment of electron-electron and electron-impurity interactions", *Superlattices and Microstructures*, vol. 27, no. 2–3, pp. 147–157, Feb. 2000.
92. A. Asenov, "Random dopant induced threshold voltage lowering and fuctuations in sub 0.1μm MOSFETs: A 3-D 'atomistic' simulation study", *IEEE Trans. Electron Dev.*, vol. 45, no. 12, pp. 2505–2513, Dec. 1998.
93. A. Asenov and S. Saini, "Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1-μm MOSFET's with epitaxial and δ-doped channels", *IEEE Trans. Electron Dev.*, vol. 46, no. 8, pp. 1718–1724, Aug. 1999.
94. L. Greengard and V. Rokhlin, "A Fast Algorithm for Particle Simulations* 1,* 2", *J. Comput. Phys.*, vol. 135, no. 2, pp. 280–292, Aug 1997.
95. R. Beatson and L. Greengard, "A short course on fast multipole methods,", *Wavelets, Multilevel Methods and Elliptic PDEs (Leicester, 1996), ser. Numer. Math. Sci. Comput. New York: Oxford Univ. Press*, pp. 1–37, 1997.
96. H. Cheng, L. Greengard, and V. Rokhlin, "A fast adaptive multipole algorithm in three dimensions", *J. Comput. Phys.*, vol. 155, no. 2, pp. 468–498, Aug. 1999.
97. FMMPART3D user's guide, version 1.0 ed., *MadMax Optics*, Hamden, CT, USA.
98. R. W. Hockney and J. W. Eastwood, "Computer Simulation Using Particles", *New York, McGraw-Hill*, 1981.
99. C. J. Wordelman and U. Ravaioli, "Integration of a particle-particle-particle-mesh algorithm with the ensemble Monte Carlo method for the simulation of ultra-small semiconductor devices", *IEEE Tran. Electron Devices*, vol. 47, no. 2, pp. 410–416, Feb. 2000.
100. W. J. Gross, D. Vasileska, and D. K. Ferry, "Ultrasmall MOSFETs: the importance of the full Coulomb interaction on device characteristics", *IEEE Electron Devices*, vol. 47, no. 10, pp. 1831–1837, Oct. 2000.
101. Allen, D. Holberg, "CMOS Analog Circuit Design", *Saunders College Publishing, New York*, 1987.
102. Bohr, Y., A. El-Mansy, "Technology for advanced high-performance microprocessors",*IEEE Trans. Electron Dev.*, vol. 45, no. 3, pp. 620–625, Mar. 1998.
103. SIA Technology Roadmap of Semiconductors: http://www.itrs.net/
104. E. H. Nicollian and A. Goetzberger, "The Si-SiO2 interface-electrical properties as determined by the metal-insulator-silicon conductance technique", *Bell Syst. Techn. J.*, vol. 46, no. 6, pp. 1055–1133, 1967.
105. J. T. Horstmann, U. Hilleringmann and K. F. Goser, "Matching analysis of deposition defined 50-nm MOSFET's", *IEEE Trans. Electron Devices*, vol. 45, no. 1, pp. 299–306, Jan. 1998.
106. P. A. Stolk, F. P. Widdershoven and D. B. M. Klaassen, "Modeling statistical dopant fluctuations in MOS transistors", *IEEE Trans. Electron Devices*, vol. 45, no. 1, pp. 1960–1971, Sep. 1998.
107. W. J. Gross, D. Vasileska and D. K. Ferry, "Three-dimensional simulations of ultrasmall metal–oxide–semiconductor field-effect transistors: The role of the discrete impurities on the device terminal characteristics", Journal of Applied Physics, vol. 91, no. 6, pp. 3737–3740, Mar. 2002.
108. H. Majima, H. Ishikuro, and T. Hiramoto, "Experimental evidence for quantum mechanical narrow channel effect in ultra-narrow MOSFET's", *IEEE Electron Dev. Lett.*, vol. 21, no. 8, pp. 396–398, Aug. 2000.

109. H.R. Khan, D. Vasileska, S.S. Ahmed, C. Ringhofer and C. Heitzinger, “Modeling of FinFETs: 3D MC Simulation Using FMM and Unintentional Doping Effects on Device Operation”, *Journal of Computational Electronics*, Vol. 3, Nos. 3–4, pp. 337–340 (2005).
110. L. Chang, S. Tang, T.-J. King, J. Bokor, and C. Hu,“Gate length scaling and threshold voltage control of double-gate MOSFETs,” *IEDM Tech. Dig.*, vol. 6, pp. 719–722, Dec. 2000.
111. Clemens Heitzinger, Christian Ringhofer, Shaikh Ahmed and Dragica Vasileska, “3D Monte-Carlo device simulations using an effective quantum potential including electron-electron interactions”, DOI 10.1007/s10825-006-0058-x, *Journal of Computational Electronics*, Volume 6, Numbers 1—3/September, pp. 15—18, 2007.

http://www.springer.com/978-1-4419-8839-3

Nano-Electronic Devices
Semiclassical and Quantum Transport Modeling
Vasileska, D.; Goodnick, S. (Eds.)
2011, XI, 441 p., Hardcover
ISBN: 978-1-4419-8839-3