

Integration of the Density Gradient Model into a General Purpose Device Simulator

ANDREAS WETTSTEIN^{a,*}, OLEG PENZIN^b and EUGENY LYUMKIS^b

^aISE Integrated Systems Engineering AG, Balgriststrasse 102, 8008 Zürich, Switzerland; ^bISE Integrated Systems Engineering, Inc., 111 N. Market Street, San Jose, CA 95113, USA

(Received 18 June 2001)

A generalized Density Gradient model has been implemented into the device simulator Dessis [DESSIS 7.0 reference manual (2001). ISE Integrated Systems Engineering AG, Balgriststrasse 102, CH-8008 Zürich]. We describe the multidimensional discretization scheme used and discuss our modifications to the standard Density Gradient model. The evaluation of the model shows good agreement to results obtained by the Schrödinger equation.

Keywords: Density gradient model; Quantization; Discretization; Device simulation; Resonant tunneling

INTRODUCTION

Today's microelectronic devices are so small that quantum mechanical effects are important. Nonetheless, industrial device simulators are based on the drift–diffusion and hydrodynamical equations, which are classical. While it is possible to couple the Schrödinger equation with classical device simulators [2,3], the numerical complexity and the nonlocality of the former increases computation time and often leads to convergence problems. On the other hand, heuristic models like the van Dort model [4] allow to predict terminal characteristics with sufficient accuracy, but do not model device internal quantities like the density distribution correctly. Hence, one of the big advantages of device simulation, the possibility to take a look inside the device, is compromised. Furthermore, the van Dort model was derived with bulk MOSFETs in mind and does not describe geometric quantization. Therefore, it cannot be applied to ultrathin SOI MOSFETs or quantum wells.

The Density Gradient model and similar models [5–9] take an intermediate position between heuristic and full quantum mechanical descriptions. They are able to predict both the terminal characteristics and the density distribution in the device. They are formulated in terms

of partial differential equations and therefore tractable with the numerical methods commonly used for classical device simulation.

While the Density Gradient model has been around in the scientific literature for many years, efficient discretization schemes are still an ongoing research topic [10,11]. Our implementation of the Density Gradient model aims at providing a useful tool for industrial users of device simulation. To meet this goal, numerical robustness, efficiency, accuracy, and ease of use have high priority, which demands a high quality discretization scheme. Also, compatibility with other transport models implemented in the simulator and restrictions due to the available meshes have to be kept in mind.

In this paper, we describe a discretization scheme for the Density Gradient Model suitable for multidimensional devices which fulfills all the aforementioned requirements to be used in day-to-day industrial device simulation. We introduce a model modification which allows to use of the Density Gradient model in parts of the device where the density is unknown. This contributes to the speed and stability, and, most importantly, to the ease of use of our implementation. We present a thorough evaluation of the Density Gradient model, identifying its abilities and limitations.

*Corresponding author. Tel.: +41-1-389-97-17. E-mail: wettstae@ise.ch

THE DENSITY GRADIENT MODEL

General Approach for Quantum Effects in a Classical Framework

To include quantization effects into a classical device simulator, we introduce an additional potential like quantity Λ into the current equation (omitting transient terms)

$$\vec{e}\vec{j} = -\mu\nabla nk_B T - \mu n \nabla (\Phi + \Phi_m + \Lambda), \quad (1)$$

where \vec{j} is the particle current density, n is the particle density, T the carrier temperature, μ the mobility, $\Phi = q\phi + \Phi_B$ the potential (consisting of both a electrostatic part ϕ and a band edge contribution Φ_B), and $\Phi_m = -(3/2)k_B T \log[m/m_0 \cdot T/T_0]$ (with the local density of states effective mass m and arbitrary normalization constants m_0 and T_0) a term related to the driving force that stems from the variation of the band edge density of states. k_B is the Boltzmann constant, e is the elementary charge, and q the particle charge. The signs of the potential like terms (Λ , Φ , Φ_m , $q\phi$, and Φ_B) are chosen such that increasing their value decreases the carrier density.

The “classical” part of the device equation consists of Eq. (1) together with the continuity equations for electrons and holes, and the Poisson equation,

$$\frac{\partial n}{\partial t} + \nabla \cdot \vec{j} = G - R, \quad (2)$$

$$\nabla \epsilon \cdot \nabla \phi = -\rho \quad (3)$$

where G and R are the generation and recombination rates, and ρ is the total space charge density (which depends on n). To keep the notation simple, we do not write down separate equations for electrons and holes in this paper.

The quantum mechanical part of the simulation is contained entirely in Λ . This quantity can be obtained by various means, according to the specific quantization model chosen. For example, in the popular van Dort model [4], Λ is a simple function of the local electric field, and in a Schrödinger equation based approach, it is given in terms of wave functions and eigenenergies and thus depends nonlocally on the potential. In the case of the Density Gradient model, Λ is the solution of a partial differential equation.

Partial Differential Equation for Λ

For the Density Gradient model [5], Λ is given by

$$\begin{aligned} \Lambda &= -\frac{\gamma \hbar^2}{12m} \left\{ \nabla^2 \log n + \frac{1}{2} (\nabla \log n)^2 \right\} \\ &= -\frac{\gamma \hbar^2}{6m} \frac{\nabla^2 \sqrt{n}}{\sqrt{n}}, \end{aligned} \quad (4)$$

where γ is a fitting parameter to be discussed below. Equations (1), (2), and (4) together are equivalent to

formulations of the Density Gradient model found in the literature [6,8], where the current continuity equation commonly contains density derivatives up to fourth order. Splitting this equation into two second-order equations has several advantages. The computation of the quantum effects described by Λ is cleanly separated out into it's own equation, making it easy to activate or to deactivate quantum effects according to user specification. The discretization method for the continuity Eq. (2) needs not to be changed; in particular, the stencil necessary to discretize it needs not to be increased. The equation for Λ is also second order and can also be discretized with the same stencil (see “Discretization Scheme” Section for a description of the discretization).

According to Eq. (4), Λ depends nonlinearly on the density n . On the other hand, by Eq. (1), the current density \vec{j} and hence the density depend on Λ and the electrostatic potential ϕ . Hence, Eqs. (2)–(4) form a nonlinear system of differential equations which have to be solved self-consistently; if both electrons and holes are important the number of coupled equations raises to five.

Defining the quasi Fermi potential E_F , we can rewrite the density as

$$n = n_0 \exp \left[\frac{E_F - \Phi - \Phi_m - \Lambda}{k_B T} \right], \quad (5)$$

where n_0 is a constant prefactor. In terms of the “smoothed potential” $\bar{\Phi} = \Phi + \Phi_m + \Lambda$, Eq. (5) allows to reformulate Eq. (4) as

$$\Lambda = -\frac{\hbar^2 \gamma}{12m} \left\{ \nabla \cdot \nabla \frac{E_F - \bar{\Phi}}{k_B T} + \frac{1}{2} \left(\nabla \frac{E_F - \bar{\Phi}}{k_B T} \right)^2 \right\}. \quad (6)$$

This formulation is equivalent to Eq. (4) but more convenient for implementation (see Eq. (7) in “Discretization Scheme” Section) and easier to generalize (see “Model Modification of Oxides” Section).

Boundary Conditions

For the new differential equation for Λ , additional boundary conditions are needed. At ohmic contacts, Dirichlet boundary conditions are used, $\Lambda = 0$. For electron and hole equations, equilibrium boundary conditions at ohmic contacts (i.e. the quasi Fermi potentials equal the applied voltage) are imposed. The boundary value of the electrostatic potential at ohmic contacts is determined from the condition of charge neutrality. In short, as $\Lambda = 0$, ohmic contacts are treated like in classical simulations.

At all other boundaries, we use homogeneous Neumann boundary conditions for Λ , i.e. $\partial_{\hat{n}}(E_F - \bar{\Phi})/k_B T = 0$, where $\partial_{\hat{n}}$ denotes the normal derivative on the boundary. Apart from Schottky contacts, also metal gate contacts

(directly on top of gate oxide) and open (noncontact) boundaries are handled this way.

At open boundaries, according to usual “classical” boundary conditions for the current and the electrostatic potential (which we keep using), the normal derivatives of the quasi Fermi potential and the electrostatic potential vanish (i.e. there is no current or field flux through the boundary), $\partial_{\hat{n}}E_F = \partial_{\hat{n}}\phi = 0$. Hence, $\partial_{\hat{n}}\Lambda = -\partial_{\hat{n}}(\Phi_B + \Phi_m)$. Typically, the band structure does not change close to the boundary, and therefore the right hand side of this equation also vanishes, that is, the Neumann boundary condition reduces to $\partial_{\hat{n}}\Lambda = 0$. Like the homogeneous boundary condition for the electrostatic potential, for some parts of the boundary (typically, at the sides of the device), this can be justified by assuming devices identical to the simulated one next of the device. For other parts of the boundary, the boundary conditions remain arbitrary.

It must be stressed that the question of boundary conditions has received little attention in the literature and we are in fact not aware of any work that dedicates a sound physical discussion to the matter. This is surprising, as the model is not completely specified without giving boundary conditions. However, Λ rapidly drops to zero away from big potential fluctuations, and therefore the boundary conditions for Λ are not important if they are imposed far (compared to $\sqrt{\hbar^2/2mk_B T}$) away from the interesting parts of the device (like the channel in MOSFETs). Nonetheless, for some applications like proper description of Schottky contacts, a more thorough investigation of the problem remains of practical interest.

Continuity Conditions

At interfaces of different materials, jumps in the band edge energy Φ_B and the effective mass m can occur. We require Λ to remain finite at the interfaces. As E_F is continuous, this implies that $\tilde{\Phi}$ and $\partial_{\hat{n}}(E_F - \tilde{\Phi})/k_B T$ are also continuous at material interfaces. The former condition simply means that the density changes continuously across the interface, a property which physically is a direct consequence of the continuity of the wave functions.

DISCRETIZATION SCHEME

While the Density Gradient method has been used in the scientific literature for some time, most authors do not report on their numerical approach at all [9,7,12]. Some papers report a finite difference scheme [6,8], and in two works a finite element scheme is used [13,14]. Recently, a nonlinear discretization was described in Ref. [10], which is reported to be far superior to more straightforward schemes linear in \sqrt{n} .

Our own discretization scheme is based on a box method (finite volume) approach using a nonlinear

discretization that evolved from the discretizations investigated in Ref. [15]. We discretize Eq. (6) as

$$\frac{6m_i}{\hbar^2}\Lambda_i \sum_{e_i} \frac{\Omega_{e_i}}{\gamma_{e_i}} = \sum_{j,e} \begin{cases} \frac{\sigma_{ij,e}}{l_{ij}} (1 - \exp[\frac{1}{2}\{\tilde{\Phi}_i - \tilde{\Phi}_j\}]) & \text{for } \tilde{\Phi}_i < \tilde{\Phi}_j \\ \frac{\sigma_{ij,e}}{2l_{ij}} \{\tilde{\Phi}_j - \tilde{\Phi}_i - \frac{1}{4}(\tilde{\Phi}_j - \tilde{\Phi}_i)^2\} & \text{otherwise} \end{cases} \quad (7)$$

with the abbreviation $\tilde{\Phi}_i = (\Phi_i - E_{F,i})/k_B T_i$. The j summation goes over all vertices neighboring vertex i , and the e sum is on all elements adjacent to the edge that connects vertices i and j . The e_i sum is on all elements vertex i belongs to. Ω_{e_i} is the volume of the fraction of box i in element e_i , l_{ij} is distance of vertices i and j , and $\sigma_{ij,e}$ the area of that part of the face between the vertices that is in element e . Otherwise, indices i and j indicate discretized values of the respective variables on these vertices.

While this particular discretization scheme was picked because in numerical experiments it proved to give superior convergence behavior when compared to a couple of alternatives [11,15], there are a few features which hint at its quality. Most importantly, the right hand side of Eq. (7) gives an M-matrix for the Jacobian with respect to $\tilde{\Phi}$. This highly desirable feature could be obtained using the exponential expression in the upper line of Eq. (7) always. However, for $\tilde{\Phi}_{i,e} > \tilde{\Phi}_{j,e}$, the exponential function easily leads to numerical overflow in the beginning of the Newton iteration, far from convergence. On the other hand, for $\tilde{\Phi}_{i,e} < \tilde{\Phi}_{j,e}$, the expression in the lower line does not guarantee the M-matrix property. Using the upwinding scheme (7) the desirable features of both schemes are combined.

To compare Eq. (7) to other discretizations found in the literature, we could reexpress our discretization scheme in terms of \sqrt{n} . The upper line in Eq. (7) corresponds to a “linear” discretization in \sqrt{n} . The lower line is a “nonlinear” discretization, but somewhat different from that described in Ref. [10]. To our experience, the discretization scheme (7) gives sufficient numerical accuracy using similar grid spacings than for classical device simulations.

As the expression for the current (1) keeps the same form as for the classical drift–diffusion model, we are able to discretize the continuity equations (2) using the Scharfetter–Gummel scheme [16] which is known to be highly accurate. The discretization of the Poisson equation follows straightforward from the box method, assuming the potential ϕ changes linearly between neighboring grid points. The discretization of all equations involves nearest neighbor information only and can be done efficiently using only data structures (geometrical coefficients) already available for the implementation of the classical equations.

MODEL MODIFICATION FOR OXIDES

The electron and hole density in insulators is small and therefore often unimportant. For example, the density in thick buried oxides of SOI devices or in the spacers between gate and source and drain in MOSFETs is mostly irrelevant. If one does not care for tunneling, even the density in the gate oxide is not important. Therefore, DESSIS does not solve the continuity equations in insulators, which increases both speed and robustness of the simulation. However, in order to use the original Density Gradient model according to Eq. (6) the density n has to be known everywhere in the device. There are two solutions to this problem: first, one can replace the insulators by wide band gap semiconductors. Then, the density is computed in the entire device and Eq. (6) can be used. The price to pay is that the users are forced to modify their structure, and computation time is increased. The second solution is to modify the model so that it works without explicit knowledge of the density.

In order to obtain an approximation for Λ which does not depend on the density n , we generalize the expression for $\tilde{\Phi}$ to $\tilde{\Phi} = (\tilde{\Phi} + (\eta - 1)q\phi - \xi E_F)/k_B T$. Two new coefficients, ξ and η , have been introduced. In the original model, $\xi = \eta = 1$. These are the values we continue to use in semiconductors. However, in insulators, we use $\xi = \eta = 0$. By using $\xi = 0$, Λ in insulators is expressed in terms of the band structure alone. In fact, in the derivation found in Ref. [8], first an expression for Λ which does not depend on E_F is derived, and only then the Fermi potential comes in by an *ad hoc* modification of this expression. Regarding this derivation, the model with $\xi = 0$ is as well justified as the original model with $\xi = 1$.

In contrast to ξ , the introduction of the parameter η and the choice of the value $\eta = 0$ has no theoretical justification and was motivated by numerical experiments only. In Fig. 1, it is shown that for $\xi = 0$, $\eta = 1$, at high gate voltages, the capacity–voltage curves lift off, and the drain current drops rapidly. The original Density Gradient model ($\xi = \eta = 1$) does not show this strange behavior. When we introduce the additional parameter η and put it to zero wherever ξ is zero (i. e. in the oxide), the artifacts disappear as well (see curve for $\xi = \eta = 0$ in Fig. 1), and the results are practically identical to those obtained by the original model.

To summarize these findings, we have three model variants each of which has one disadvantage:

- The original model ($\xi = \eta = 1$ everywhere) which works fine, but forces us to compute the density in the barrier (by treating the barrier as a wide band gap semiconductor).
- A physically motivated modification ($\xi = 0$, $\eta = 1$ in insulators) which does not require the density in the barrier, but does not work correctly for high gate voltages.

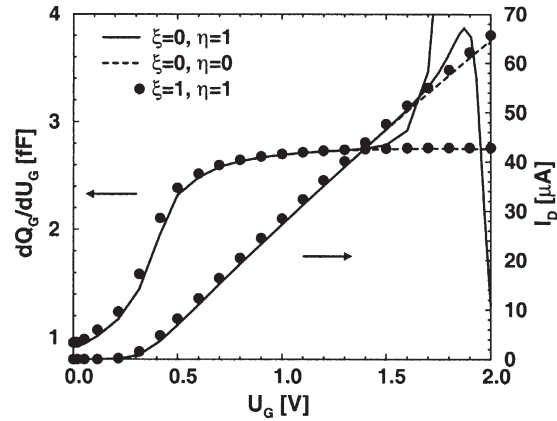


FIGURE 1 Gate capacity $\partial Q_G/\partial U_G$ and drain current I_D as a function of gate voltage U_G for a nMOSFET with gate length 300 nm, gate width 1 μm , and gate oxide thickness 3 nm for drain voltage $U_D = 10\text{ mV}$.

- A modified model ($\xi = \eta = 0$ in insulators) which does not need the density in the barrier and gives reasonable results even for high gate voltages, but lacks theoretical justification.

For practical purposes, only $\xi = \eta = 1$ and $\xi = \eta = 0$ are acceptable.

The artifacts observed in Fig. 1 are interesting from the standpoint of fundamental theory, as they clearly favor the standard formulations of the Density Gradient model which depend on the quasi Fermi energy (i.e. $\xi = 1$) over formulations which do not ($\xi = 0$). With regard to the microscopic derivation of the model (see, for example, Refs. [15,17]), this is not obvious, as the derivations assume a near equilibrium situation where spatial variations of the quasi Fermi energy are negligible.

How does the strange behavior for $\xi = 0$, $\eta = 1$ arise? The quasi Fermi potential drops over the barrier. This drop partially cancels the drop of the band edge due to the applied voltage (see Fig. 2, right plot). When $\xi = 0$ and $\eta = 1$, no such compensation takes place. If the applied voltage is high, the barrier between the channel and the gate is just a small triangle. Where the artifacts for $\xi = 0$, $\eta = 1$ appear, we observe that the small barrier has been completely smoothed away by the quantum potential (Fig. 2, left).

By using $\eta = 0$, the reduction of the barrier by the applied electric field is neglected, which partially mimics the compensating behavior of the quasi Fermi potential. The quasi Fermi potential and the band edge do not behave exactly the same (see Fig. 2, right plot), and therefore, using $\eta = \xi = 0$ is not exactly equivalent to using $\eta = \xi = 1$. Nonetheless, the effect is very similar, which strengthens our confidence in the otherwise completely heuristic model variant $\eta = \xi = 0$.

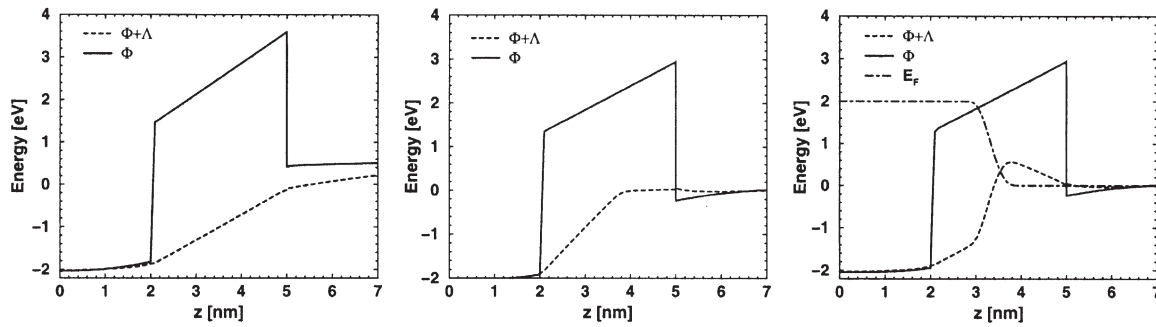


FIGURE 2 Potential Φ , $\Phi + \Lambda$, and quasi Fermi energy along a cut perpendicular to the silicon-oxide interface in mid-channel of the device used for Fig. 1, at gate voltage $U_G = 2$ V. Left plot: $\xi = 0$, $\eta = 1$ in oxide. Middle: $\xi = \eta = 0$ in oxide. Right plot: $\xi = \eta = 1$ in oxide.

MODEL EVALUATION

Comparison to Schrödinger Results

For MOS diodes or MOSFETs with moderate channel length, the 1D Schrödinger equation can be considered as a description of quantization effects which is exact with respect to the underlying theoretical framework (effective mass approximation). Our goal in this section is to establish the validity of the Density Gradient method within this framework. We do so by comparing the results obtained by this model to results obtained by the Schrödinger equation solver available in DESSIS [1]. For this theory-to-theory comparison it is sufficient to consider near equilibrium bias conditions and to ignore secondary effects such as modified mobilities or recombination rates.

In the oxide, we have chosen the fit factor $\gamma = 1$. While this value determines the carrier concentration at the silicon-oxide interface, it has little influence on the terminal characteristics. To fit the Schrödinger results, in silicon, $\gamma = 3.6$ has been used. This value can be understood as follows. The conduction band of silicon has six equivalent minima, in each of which the band structure can be described in terms of an anisotropic effective mass. The lowest eigenenergy of the Schrödinger equation is

determined by the largest effective mass component m_z of the effective mass tensor (at least for the $\langle 100 \rangle$ substrate orientation, which was used in our simulations and which is typically present in actual devices). For $T = 300$ K, and the parameters used in our simulation, the ratio of the density of states effective mass m (which enters the Density Gradient model) and m_z is about 1.2. Furthermore, it is known from the literature [18] that by replacing the 12 in the denominator of Eq. (4) by 4, one arrives at an equation that can be derived directly from the Schrödinger equation and can be considered as the low temperature limit of the Density Gradient model. Hence, the value $\gamma = 3.6$ is caused by the fact that m_z rather than m is the mass governing quantization effects, and that the potential variations at the silicon-oxide interface are large compared to the thermal energy (3.15 eV vs. 25 meV), and therefore the low temperature limit of the model has to be picked.

Given an explanation for the value of the fit factor, it is not surprising that it is independent from device characteristics like doping concentration and oxide thickness. This is illustrated in Fig. 3 by a comparison of capacity–voltage curves for a range of channel doping concentrations and oxide thicknesses.

One of the advantages of the Density Gradient model over heuristic quantization models is that it correctly

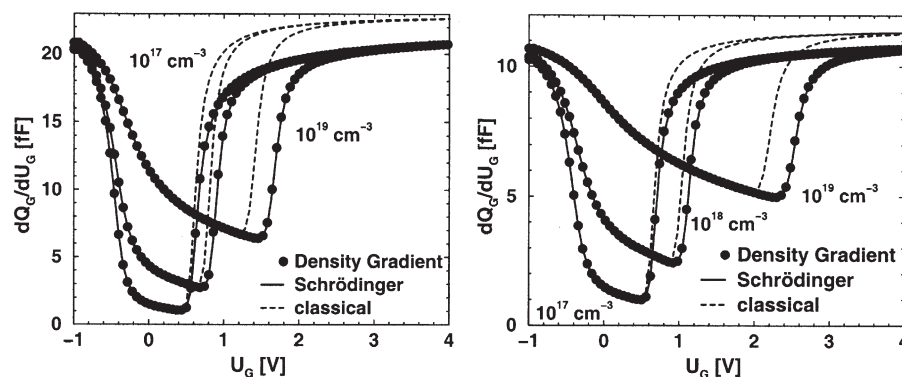


FIGURE 3 Capacity–voltage characteristics for a MOS diode with metal gate, 1.5 nm (left) and 3 nm (right) oxide thickness and area of $1 \mu\text{m}^2$. We compare the results of the Density Gradient method to those of classical simulations and simulations with the Schrödinger equation. For each oxide thickness, curves for substrate dopings of 10^{17} cm^{-3} , 10^{18} cm^{-3} , and 10^{19} cm^{-3} are shown. All curves were obtained with $\gamma = 3.6$ in silicon and $\gamma = 1$ in oxide.

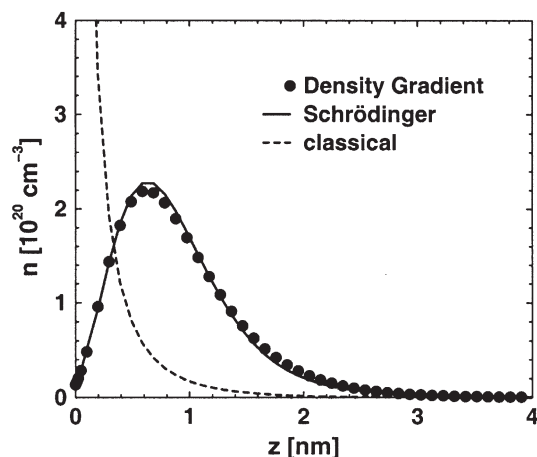


FIGURE 4 Electron density in channel of a MOS diode with 3 nm oxide thickness and 10^{18} cm^{-3} doping concentration for a gate voltage of 5 V. The silicon-oxide interface is located at $z = 0$.

predicts the density distribution in the channel, see Fig. 4.

The Density Gradient method can be applied to other structures than bulk MOSFETs, too. For illustration, we have performed simulations of a double gated SOI MOSFET with oppositely doped gates, similar to the structure suggested in Ref. [19]. The opposite doping of the gates leads to a builtin potential between them, and hence generates an electric field across the channel. The density profile in the channel is shown in Fig. 5, and is in good agreement to the density obtained by the Schrödinger equation. The gate voltage dependence of the drain current is also correctly predicted. In Fig. 6, we show calculations for silicon layer thickness from 10 nm down to 2.5 nm. The Density Gradient method reproduces the Schrödinger results for all layer thicknesses.

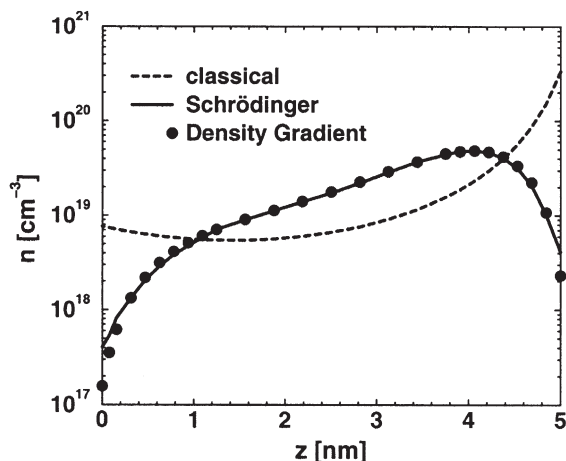


FIGURE 5 Electron density along a cut perpendicular to the silicon-oxide interfaces in mid-channel of a double gated SOI MOSFET with 80 nm channel length and 1.5 nm gate oxide thickness. The plot was obtained for gate voltage $U_G = 1 \text{ V}$ and drain voltage $U_D = 50 \text{ mV}$. The gate to the left side is assumed p -doped, the gate on the right as n -doped, which was modeled assuming appropriate work functions for the contacts.

The difference between the curves obtained with the classical simulation and the quantum mechanical simulations gets bigger with decreasing silicon layer thickness, reflecting the increased quantization when squeezing the electrons into a narrower well. The Density Gradient method correctly describes this geometrical quantization.

Quantization in Poly Gates

When compared to classical calculations, quantization reduces the density of carriers in a semiconductor close to interfaces to insulating regions. It is not necessary that the carriers are confined in a potential minimum like in an inverted MOSFET channel—the presence of the potential barrier the insulator represents is sufficient to “repel” the carriers. Thus, even for flatband conditions (or for majority carrier depletion), a quantization effect for the majorities remains (see Fig. 7). If the classical majority carrier density is big, i.e. from accumulation to moderate depletion, the modification of the charge density can be quite substantial. This is important for poly gates, where the quantum mechanical density reduction near the interface gives rise to a space charge generated by the ionized dopants, assuming the latter represent a homogeneous charge density. This space charge attracts minority carriers to the MOSFET channel and can compensate the threshold voltage shift due to quantization in the channel itself [20].

In Fig. 8, we compare a capacity–voltage curve obtained by a classical simulation to curves obtained by the Density Gradient model. In one case, the Density Gradient model is used everywhere, in the other case, it is used only in silicon (Eq. (6) is solved globally, but in the poly gate, Λ is discarded in Eq. (1)). For a poly doping of 10^{19} cm^{-3} , the poly quantization effect is negligible, and the two quantum mechanical simulations agree. For 10^{20} cm^{-3} , the poly quantization completely compensates the threshold voltage shift caused by the channel quantization, and also results in a further reduction of the gate capacity in inversion. The different impact of poly quantization for the different dopings is due to the smaller space charge and stronger classical depletion in the lower doped poly gate.

The same results for poly quantization can also be obtained by the Schrödinger equation available in DESSIS [11], and therefore it is clear that this effect is not an artifact of the Density Gradient model. However, there are theoretical and experimental problems with the poly quantization effect. Theoretically, it is not clear whether it is possible to describe the poly similar to normal silicon, given the huge amount of traps and the band tails due to the high disorder. The disorder might render the effective mass approximation invalid, which is the base for the Density Gradient method (and the Schrödinger equation as implemented in DESSIS). Experimentally, the poly doping concentration at the interface is not available from direct

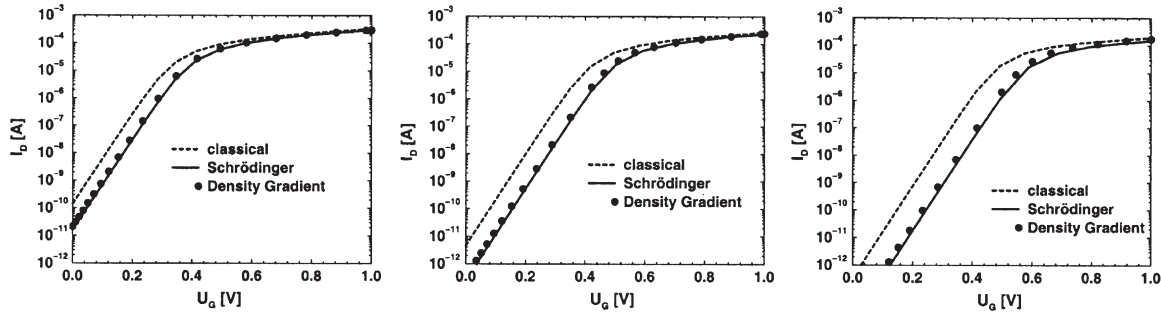


FIGURE 6 Drain current vs. gate voltage of the SOI MOSFET with 10 nm (left), 5 nm (middle) and 2.5 nm (right) silicon layer thickness.

SIMS measurements, and it is therefore difficult to separate effects of doping distribution and quantum mechanical depletion in poly silicon.

Tunneling

One thing which made the Density Gradient and similar models popular is their ability to describe resonant tunneling. In Fig. 9, we show an example of such a computation using a silicon-oxide resonant tunneling diode. To obtain this example, the current Eq. (1) has to be solved in the barriers, that is, the oxide is treated as a wide band gap semiconductor with $\xi = \eta = 1$ (see the discussion in “Model Modification for Oxides” Section).

The Density Gradient method handles tunneling in a manner completely different from conventional models employed in device simulators (see, for example, [21]). In the latter, carriers disappear on one side of the barrier and reappear on the other side, which means that tunneling is treated like a nonlocal recombination–generation mechanism. In contrast, the Density Gradient model describes the tunneling by a normal drift–diffusion current which

flows though a barrier strongly reduced by Λ (this barrier reduction is shown in Fig. 2).

Some doubt regarding the theoretical validity of the Density Gradient approach to tunneling remains. To illustrate these concerns, we have plotted the density distribution in the barrier (Fig. 10). The density decays into the barrier from both sides. The tunneling current is limited by the lowest value of the density in the barrier, which is reached somewhere in the middle of the barrier. Hence, the tunneling current is mainly determined by the decay of the density over about *half* of the barrier. However, physically, the tunneling current is determined by the decay of wave functions over the *entire* barrier. This problem has been addressed in the literature by extending the plain Density Gradient model to use two instead of one new variable in the barrier, corresponding to carriers tunneling in backward and forward direction [22,23]. Unfortunately, this extension is one dimensional by its very formulation (the concept of “backward” and “forward” has no immediate generalization to higher dimensions). For multi dimensional device simulation, no solution to this theoretical problem is currently known.

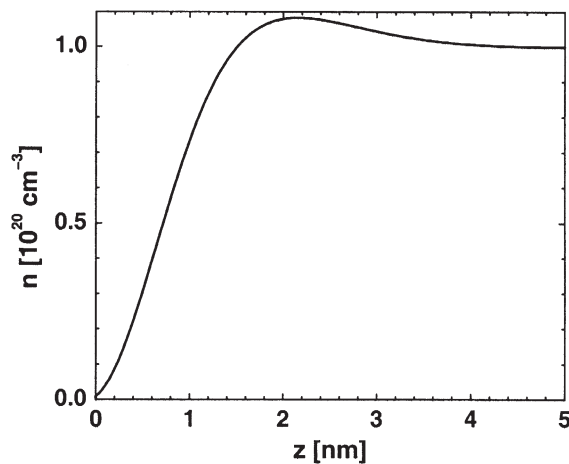


FIGURE 7 Density gradient result for the electron density in a 10^{20} cm^{-3} *n*-doped poly gate of a MOS diode with 3 nm oxide thickness and $5 \times 10^{17} \text{ cm}^{-3}$ substrate *p*-doping, for 1 V gate voltage. The oxide–poly interface is located at $z = 0$.

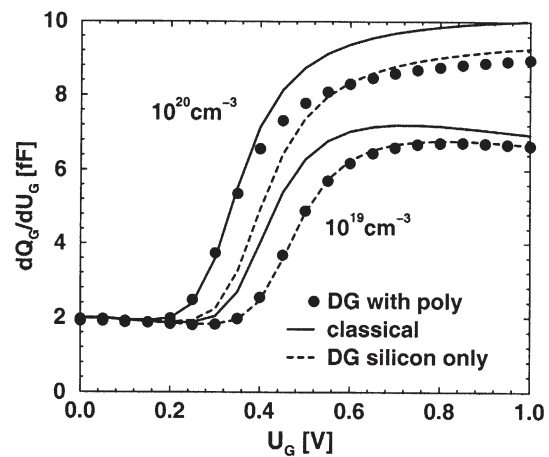


FIGURE 8 Low frequency capacitance–voltage plot for NMOS diode with 3 nm oxide thickness, $5 \times 10^{17} \text{ cm}^{-3}$ doping concentration, and poly dopings of 10^{19} cm^{-3} and 10^{20} cm^{-3} , obtained by a classical simulation, a calculation with the Density Gradient model everywhere in the device, and with the Density Gradient model correction in silicon only (but not in poly).

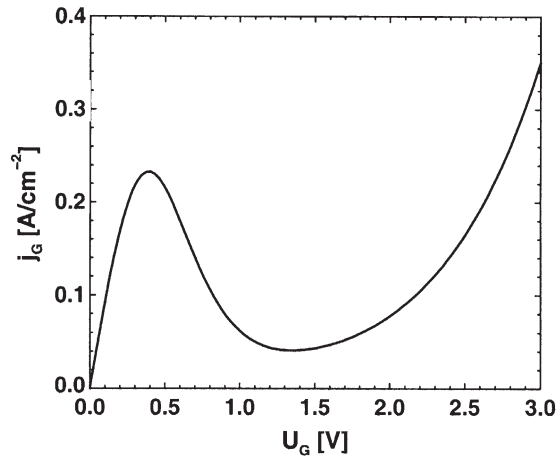


FIGURE 9 Current density as a function of applied voltage through resonant tunneling diode formed by two 1 nm oxide barriers separated by a 5 nm silicon layer.

From a pragmatic point of view, it might well be possible to use the plain Density Gradient model to describe tunneling if one makes use of the available fitting parameters. Even physically more accurate tunneling models based on the Schrödinger equation usually require an adjustment of fit parameters, and the effective mass approximation they employ becomes questionable for thin oxides. In Fig. 11, we compare tunneling currents obtained by Density Gradient method to currents computed with a Schrödinger equation based tunneling model [24]. To fit the former to the latter, the mobility in the oxide has been fitted to $\mu = 0.05 \text{ cm}^2/\text{Vs}$. All other parameters kept their default values. The actual fit was done for 2 nm oxide thickness. The figure shows that the Density Gradient method captures the thickness dependence of the tunneling current reasonably well.

SUMMARY

We described a robust discretization scheme for the Density Gradient model which is applicable for multi

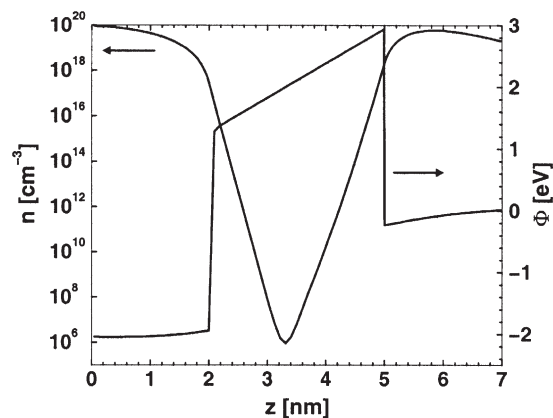


FIGURE 10 Density distribution and potential profile in the 3 nm MOSFET structure at $U_G = 2 \text{ V}$.

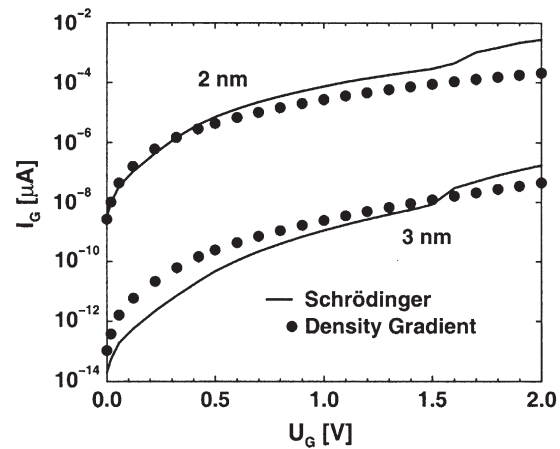


FIGURE 11 Gate tunneling currents computed with the Density Gradient method and a Schrödinger equation based tunneling model for nMOSFETs with gate length 300 nm, gate width $1 \mu\text{m}$, and gate oxide thickness of 2 and 3 nm.

dimensional devices and leads to good convergence behavior. By a minor modification of the original model, we are able to use the method in insulators where the density is not computed.

While carrier quantization in the channel, including geometrical quantization, is well described by the Density Gradient model, its ability to describe tunneling is unclear, and the need for future research remains.

The results of the Density Gradient method are demonstrated to be in excellent agreement to results obtained by the Schrödinger equation, and the value of the only fit factor in the model can be understood on a physical base. The Density Gradient model is hence demonstrated to combine high physical accuracy with the numerical stability required by professional users of TCAD software.

Acknowledgements

The authors wish to thank Dr Boris Polsky (ISE, Inc.) for useful discussions and his critical examination of our implementation.

References

- [1] DESSIS 7.0 reference manual (2001). ISE Integrated Systems Engineering AG, Balgriststrasse 102, CH-8008 Zürich.
- [2] Stern, F. (1970) "Iteration methods for calculating self-consistent fields in semiconductor inversion layers", *J. Comput. Phys.* **6**(1), 56–67.
- [3] Wettstein, A., Schenk, A., Scholze, A., Garretón, G. and Fichtner, W. (1997) "Charge carrier quantization effects in double-gated SOI MOSFETs", *ULSI Sci. Technol.* **1997**, 613–621.
- [4] van Dort, M.J., Woerlee, P.H. and Walker, A.J. (1994) "A simple model for quantisation effects in heavily-doped silicon MOSFETs at inversion conditions", *Solid-State Electron.* **37**(3), 411–414.
- [5] Ancona, M.G. and Tiersten, H.F. (1987) "Macroscopic physics of the silicon inversion layer", *Phys. Rev. B* **35**(15), 7959–7965.
- [6] Ancona, M.G., Yu, Z., Lee, W.-C., Dutton, R.W. and Voorde, P.V. (1999) "Simulation of quantum confinement effects in ultra-thin-oxide MOS structures", *J. Technol. Comput. Aided Design* **11**.

- [7] Zhou, J.-R. and Ferry, D.K. (1992) "Ballistic phenomena in GaAs MESFETs: modelling with quantum moment equations", *Semicond. Sci. Technol.* **7**(3B), B546–B548.
- [8] Gardner, C.L. (1994) "The quantum hydrodynamic model for semiconductor devices", *SIAM J. Appl. Math.* **54**(2), 409–427.
- [9] Grubin, H.L., Govindan, T.R., Kreskovsky, J.P. and Strosio, M.A. (1993) "Transport via the Liouville equation and moments of quantum distribution functions", *Solid-State Electron.* **36**(12), 1697–1709.
- [10] Ancona, M.G. and Biegel, B.A. (2000) "Nonlinear discretization scheme for the density-gradient equations", *SISPAD 2000*, 196–199.
- [11] Wettstein, A., Schenk, A. and Fichtner, W. (2001) "Quantum device-simulation with the density-gradient model on unstructured grids", *IEEE Trans. Electron Devices* **48**(2), 279–284.
- [12] Zhou, J.-R. and Ferry, D.K. (1992) "Simulation of ultra-small GaAs MESFET's using quantum moment equations-II: velocity overshoot", *IEEE Trans. Electron Devices* **39**(8), 1793–1796.
- [13] Chen, Z., Cockburn, B., Gardner, C.L. and Jerome, J.W. (1995) "Quantum hydrodynamic simulation of hysteresis in the resonant tunneling diode", *J. Comput. Phys.* **117**(2), 274–280.
- [14] Chen, Z. (1996) "A finite element method for the quantum hydrodynamic model for semiconductor devices", *Comput. Math. Appl.* **31**(7), 17–26.
- [15] Wettstein, A., "Quantum Effects in MOS Devices" Ph.D. Thesis (Hartung-Gorre, Konstanz).
- [16] Scharfetter, D.L. and Gummel, H.K. (1969) "Large-signal analysis of a silicon read diode oscillator", *IEEE Trans. Electron Devices* **16**(1), 64–77.
- [17] Gardner, C.L. and Ringhofer, C.A. (1996) "Smooth quantum potential for the hydrodynamic model", *Phys. Rev. E* **53**(1), 157–167.
- [18] Ferry, D.K. and Zhou, J.-R. (1993) "Form of the quantum potential for use in hydrodynamic equations for semiconductor device modeling", *Phys. Rev. B* **48**(11), 7944–7950.
- [19] Suzuki, K., Tosaka, Y. and Sugii, T. (1995) "Analytical threshold voltage model for short channel n⁺-p⁺ double-gate SOI MOSFETs", *Proc. 1995 IEEE Int. SOI Conf.*, 68–69.
- [20] Spinelli, A.S., Pacelli, A. and Lacaita, A.L. (2000) "Polysilicon quantization effects on the electrical properties of MOS transistors", *IEEE Trans. Electron Devices* **47**(12), 2366–2371.
- [21] Schenk, A. and Heiser, G. (1997) "Modeling and simulation of tunneling through ultra-thin gate dielectrics", *J. Appl. Phys.* **81**(12), 7900–7908.
- [22] Ancona, M.G. (1990) "Macroscopic description of quantum-mechanical tunneling", *Phys. Rev. B* **42**(2), 1222–1233.
- [23] Ancona, M.G., Yu, Z., Dutton, R.W., Voorde, P.J.V., Cao, M. and Vook, D. (2000) "Density-gradient analysis of MOS tunneling", *IEEE Trans. Electron Devices* **47**(12), 2310–2319.
- [24] Wettstein, A., Schenk, A., Scholze, A. and Fichtner, W. (1997) "The influence of localized states on gate tunnel currents—modeling and simulation", *SISPAD 1997 Tech. Digest*, 101–104.

A. Wettstein received the Diploma in physics from the Universität Karlsruhe in 1995 and the Ph.D. degree in engineering from the Eidgenössische Technische Hochschule in Zürich in 2000. Since then is with ISE AG, Zürich, working as a software developer for device simulation.

O. Penzin received the M.S. degree in Electrical Engineering from Novosibirsk Technical University in 1984. From 1984 to 1988, he worked at Silicon Fab "Vostok" (Novosibirsk, Russia) as device engineer mainly for development of numerical models and optimization of CCD and photo detectors. From 1988 to 1996, he was researcher at Computing Center of Russian Academy of Science (Novosibirsk, Russia) and a consultant of Dawn Technologies, Inc. (Sunnyvale, CA), where his interest was in advanced semiconductor physics, Si and III–V devices, and numerical methods. Since 1996 he is with ISE, Inc. (San Jose, CA) where currently he is responsible for device simulation.

E. Lyumkis received the M.S. degree in Physics from Moscow Institute of Physics and Technology in 1972 and the Ph.D. degree in numerical mathematics from Moscow Institute of Applied Mathematics in 1977. From 1975 to 1991 he worked at the Research Institute of Applied Mathematics and Computer Science at the Latvian University, Riga, where he focused on the numerical modeling of hydrodynamical processes and crystal growth. He works on numerical simulation of semiconductor devices since 1991, formerly at Silvaco International, and since 1995 at ISE. He is presently Manager of Process and Device Development at ISE, Inc., San Jose, CA.