Efficient Multi-Dimensional Simulation of Quantum Confinement Effects in Advanced MOS Devices

Bryan A. Biegel
Mailstop T27A-2
NASA Ames Research Center
Moffett Field, CA 94035-1000
Bryan.A.Biegel@nasa.gov

Mario G. Ancona
Code 6876
Naval Research Lab
Washington, DC 20375
ancona@estd.nrl.navy.mil

Conor S. Rafferty
211 Warren St.
Noble Device Technologies
Newark, NJ 07103
conor.rafferty@nobledevice.com

Zhiping Yu
Institute of Microelectronics
Tsinghua University
Beiging, China
yuzhip@tsinghua.edu.cn

ABSTRACT

We investigate the density-gradient (DG) transport model for efficient multi-dimensional simulation of quantum confinement effects in advanced MOS devices. The formulation of the DG model is described as a quantum correction to the classical drift-diffusion model. Quantum confinement effects are shown to be significant in sub-100nm MOSFETs. In thin-oxide MOS capacitors, quantum effects may reduce gate capacitance by 25% or more. As a result, the inclusion of quantum effects in simulations dramatically improves the match between C-V simulations and measurements for oxide thickness down to 2 nm. Significant quantum corrections also occur in the I-V characteristics of short-channel (30 to 100 nm) n-MOSFETs, with current drive reduced by up to 70%. This effect is shown to result from reduced inversion charge due to quantum confinement of electrons in the channel. Also, subthreshold slope is degraded by 15 to 20 mV/decade with the inclusion of quantum effects via the density-gradient model, and short channel effects (in particular, drain-induced barrier lowering) are noticeably increased.

I. INTRODUCTION

The semiconductor industry is in a period of feverish advancement, with new generations of electronics technology being developed every 2 years [1]. For the industry-dominant MOSFET, gate lengths and oxide thicknesses of production devices will shrink below 50 nm and 1 nm respectively over the next decade [2]. Meanwhile, the classical drift-diffusion (DD) model of electron transport has steadfastly remained the dominant model for industrial numerical simulation of electronic devices [3], 50 years after the model's first description [4], and 35 years after Gummel [5] described a robust numerical solution method. This uncommon longevity stems partly from momentum (due to familiarity and experience), but it derives mainly from the fact that, with tuning for a given technology, the DD model continues to provide a combination of computational speed, robustness, and acceptable accuracy which has been difficult to match with other models. Nevertheless, there is a growing realization that the electronics industry can not finesse quantum effects with the DD model much longer [1]. In this paper, we describe a transport model which meets the challenge of combining computational efficiency, robustness, and accuracy in a multi-dimensions including quantum effects, and we present promising simulation results for this model.

Two approaches have been taken to develop models which meet this challenge: (i) employing full quantum models for conventional device simulation, and (ii) adding quantum corrections to classical models such as DD. The first approach includes the use of quantum models such as non-equilibrium Green's function [6] and Wigner function [7]. The Green's function model contains a high level of quantum mechanical and scattering detail, but in multi-dimensions it is many orders of magnitude more costly than the DD model. The Wigner function model trades quantum mechanical and scattering detail for somewhat better computational efficiency in comparison to

the Green's function model, but it suffers from unsolved numerical robustness problems [8], and is still much too costly for production-level computations, even in 2-D. We note that 1-D simulations can provide some information about electronic device operation, but 2-D (or 3-D) simulations are essential for sufficient detail and accuracy for industrial application.

The second approach for including quantum effects in device simulations is to add quantum corrections to classical models. For the DD model in a MOSFET, approximate quantum corrections have been applied to the inversion charge profile [9], the silicon band gap near the oxide [10], and both the band gap and the gate oxide thickness [11]. Alternatively, a 1-D Schrödinger computation has been added to the DD model to account for quantum confinement effects in the inversion layer [12,13]. These approaches may be computationally efficient, but they assume a particular device structure (MOSFET with planar gate) and operating region (inversion). In contrast, quantum corrections to the hydrodynamic model [14,15] are general in terms of allowed device structures and operating regions, but they suffer from significant disadvantages in computational robustness and cost versus the DD model [16].

In this work, we describe and employ a device simulation model which uses the second approach to meet the multi-dimensional quantum challenge while avoiding the disadvantages of the approaches described above. This model is the density-gradient (DG) quantum correction to the DD model [17]. DG theory has a long history, dating back to von Weizsaker's introduction [18] of gradient corrections to Thomas-Fermi theory for an inhomogeneous electron gas. Gradient corrections were carried over into density functional (DF) theory by Hohenberg and Kohn for zero temperature [19] and by Mermin for finite temperatures [20]. This approach is equivalent to the exact (non-local) many-body theory when gradient corrections to all orders are retained [19-21]. Truncated gradient expansions (as further developed by Perdew [22] and others) remain an important part of DF calculations. More recently, DF theory has been used to determine density-gradient corrections to the equation of state of the electron gas in the random phase approximation for arbitrary temperature [23]. This work is an application of DG theory to semiconductor problems. We note that DG theory is distinct from the Bohm quantum potential (which is identical in form, but has a different numerical coefficient) since DG theory applies to a many-body mixed state, while the Bohm potential applies to a single-particle pure state.

The DG model adds quantum confinement and (optionally) tunneling to the DD model in a general, compact, and computationally efficient manner. This yields a model which meets all of the requirements of a replacement for the DD model. In this work, we focus on quantum confinement effects in MOS devices, using an implementation of the DG model which does not include quantum tunneling. After describing the model, the remainder of this paper compares classical and DG model simulation results. First, we compare capacitance-voltage (C-V) curves for thin-oxide MOS capacitors, showing a significant reduction in capacitance due to quantum confinement of electrons and holes. We then show the large reduction in drive current of ultra-small MOSFETs when quantum confinement is included. We also show the degradation of small MOSFET subthreshold slope and the increase in short-channel effects (DIBL). We conclude with a summary of these results.

II. Density-Gradient Model

The drift-diffusion and density-gradient models of carrier transport in an electronic device can be written identically, including Poisson's equation and the electron and hole transport equations:

$$\nabla \cdot (\varepsilon \nabla \psi) = -\rho = -q(p - n + C); \tag{1a}$$

$$\frac{\partial n}{\partial t} = \frac{\nabla \cdot \boldsymbol{J}_n}{q} = \nabla \cdot (-n\mu_n \nabla \psi_n + D_n \nabla n); \tag{1b}$$

$$\frac{\partial p}{\partial t} = -\frac{\nabla \cdot \boldsymbol{J}_p}{q} = \nabla \cdot (p\mu_p \nabla \psi_p + D_p \nabla p); \tag{1c}$$

where ψ is electrostatic potential, n and p are electron and hole densities, C is fixed charge density, ε is permittivity, ρ is total charge density, q is electron charge, $\boldsymbol{J}$ is current density, and μ and D are mobility and diffusivity of the respective carriers.

In the classical DD model, the electron and hole "drift potentials" are just the electrostatic potential: $\psi_n = \psi_p = \psi$. In the DG model, ψ_n and ψ_p have quantum corrections:

$$\psi_n = \psi + \psi_{qn};\ \psi_{qn} \equiv 2b_n\left(\frac{\nabla^2\sqrt{n}}{\sqrt{n}}\right);\ b_n \equiv \frac{\hbar^2}{4r_n m_n q}; \tag{2a}$$

$$\psi_p = \psi + \psi_{qp};\ \psi_{qp} \equiv 2b_p\left(\frac{\nabla^2\sqrt{p}}{\sqrt{p}}\right);\ b_p \equiv \frac{\hbar^2}{4r_p m_p q}. \tag{2b}$$

The expressions for the quantum potentials ψ_{qn} and ψ_{qp} are derived from the Schrödinger equation, based on the finite curvature (energy) and strict continuity of wavefunctions [17,24]. Since the quantum potentials represent a net effect for all wavefunctions, they do not incorporate quantum mechanics exactly into the DG model, so r_n and r_p may be used as fitting parameters. In this work, we take $r_n = r_p = 3$, which is the high temperature limit [25,26]. The quantum potentials act to smooth the carrier density profiles by reducing their second derivatives (curvature). In fact, the DG model forbids discontinuities in the carrier density profiles.

We focus on quantum confinement effects in the industry-dominant device, the silicon MOSFET (Figure 1). The effects of quantum smoothing on carrier profiles in a MOSFET operating in inversion are depicted in Figure 2. Classical carrier densities (dashed curves) change abruptly at the oxide interfaces from some large external value to zero in the oxide. Quantum mechanical carrier densities (solid curves) can not change abruptly. The densities must be continuous across the oxide interfaces, resulting in significant differences in classical and quantum carrier profiles near the interfaces, and penetration of quantum carrier densities (quantum tunneling) into the oxide.

Returning to the DG model in (1) and (2), the five solution variables are ψ, n, p, ψ_{qn}, ψ_{qp}, each of which involve second-order partial differential equations (PDEs). The quantum potentials ψ_{qn} and ψ_{qp} accomplish the expected carrier profile smoothing, and so will be largest near the oxide interfaces where the classical density discontinuities need to be smoothed out. In this work, we ignore carrier densities (and thus tunneling) in the oxide, and we set the carrier densities n and p to 0 in this region. Thus, only the source-less Poisson equation (1a) is solved in the oxide, while the full DG model is solved in both the gate and substrate. In order for the carrier densities to be continuous across the oxide interfaces, they must approach zero just outside the oxide. With this and the usual boundary conditions (BCs) for the DD model, all of the BCs of the DG model are well defined except for those for the quantum potentials at the oxide interfaces.

The two quantum potentials have relatively large, unknown values at the oxide interfaces, since this is where they act to force electron and hole concentrations smoothly to zero. Thus, a suitable constraint on the values of the quantum potentials at the oxide interfaces (a Dirichlet BC)

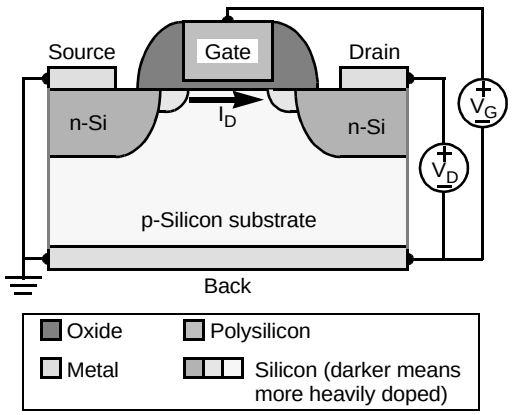


Figure 1: Basic n-MOSFET structure and biasing. Bias V_G between gate and source contacts controls current I_D flowing between source and drain.

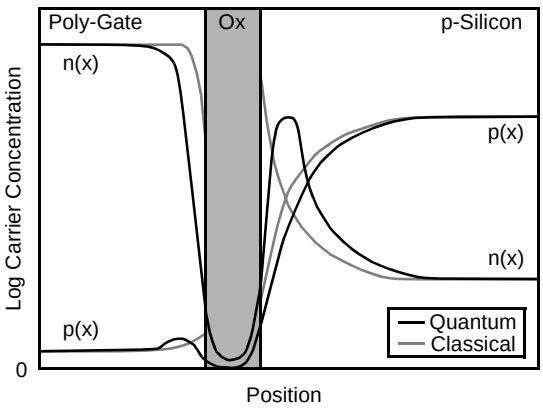


Figure 2: Schematic of classical (dashed lines) and quantum (solid lines) carrier density profiles of a vertical cut through an n-MOSFET in inversion. Classical densities are sharply peaked and discontinuous at the silicon/oxide interface, and can not penetrate into the oxide. Quantum profiles are smoothly peaked below the interface, continuous at all interfaces, and can penetrate (tunnel) into the oxide.

is not available. By the same reasoning, enforcing a Neumann BC on the (unknown) gradients of the quantum potentials is also not valid. One solution to this lack of quantum potential BCs at the oxide interfaces is to solve the entire five-PDE model in the oxide as well as in the adjoining silicon and poly gate. This also implements tunneling in the DG model. In this case, the discontinuity in the quantum potentials is determined by the silicon-oxide band offsets, and the gradients would be continuous across the interface. A model implementing this approach has been developed [28].

In order to minimize the changes from the DD to the DG model, and thus more directly compare simulation results, we have chosen in this work not to include tunneling in the DG model. Therefore, another solution to the boundary condition challenge is needed. In particular, we can use the quasi-Fermi (QF) model [29] of carrier transport, which can be described as a change of variables from the DD model. In the QF model, the continuity equations are:

$$\frac{\partial n}{\partial t} = \frac{\nabla \cdot \boldsymbol{J}_n}{q} = \nabla \cdot (-n\mu_n \nabla \phi_n); \tag{3a}$$

$$\frac{\partial n}{\partial t} = -\frac{\nabla \cdot \boldsymbol{J}_p}{q} = \nabla \cdot (p\mu_p \nabla \phi_p); \tag{3b}$$

where, including DG quantum corrections, the QF energies are:

$$\phi_n = \psi - (kT/q)\ln(n/n_i) + \psi_{qn}; \tag{4a}$$

$$\phi_p = \psi + (kT/q)\ln(p/n_i) + \psi_{qp}; \tag{4b}$$

and n_i is the intrinsic carrier concentration of the semiconductor. In (4a) and (4b), we have assumed a Maxwell-Boltzmann energy distribution of the carriers. The DG model with (more exact) Fermi-Dirac statistics has also been implemented.

Note from (3a) and (3b) that at the interface between a semiconductor and an insulator, the electron and hole QF energies ϕ_n and ϕ_p in the semiconductor have zero gradient normal to the interface, since current flow $\boldsymbol{J}_n$, $\boldsymbol{J}_p$ into the insulator is zero. [Recall that tunneling current is zero in this work.] The DG version of the QF model can use these constraints on ϕ_n and ϕ_p as the additional BCs needed to solve the quantum potential PDEs. Before writing the final DG model used for this work, we note that only steady-state simulations were used, so the time derivatives were eliminated from the continuity equations. Thus, the DG model solved in this work is:

$$\nabla \cdot (\varepsilon \nabla \psi) + q(p - n + C) = 0, \tag{5a}$$

$$\nabla \cdot (n \mu_n \nabla \phi_n) = 0, \tag{5b}$$

$$-\nabla \cdot (p \mu_p \nabla \phi_p) = 0, \tag{5c}$$

$$\sqrt{n} \psi_{qn} - 2 b_n \nabla^2 \sqrt{n} = 0, \tag{5d}$$

$$\sqrt{p} \psi_{qp} + 2 b_p \nabla^2 \sqrt{p} = 0. \tag{5e}$$

The five solution variables for these PDEs, in order, are ψ, ϕ_n, ϕ_p, n, p. Expressions for the quantum constants b_n and b_p are given in (2a) and (2b). From (4a) and (4b):

$$\psi_{qn} = \phi_n + (kT/q) \ln(n/n_i) - \psi; \tag{6a}$$

$$\psi_{qp} = \phi_p - (kT/q) \ln(p/n_i) - \psi. \tag{6b}$$

Note that the DG model in (5a) - (5e) is generic, in the sense that it can be applied to any electronic device structure in 1-D, 2-D, or 3-D. By contrast, other quantum-DD models often incorporate structure-specific and localized quantum corrections [9-13], and may not allow tunneling to be included.

For this work, we used fixed mobilities of μ_n = 1500 cm^2/Vs and μ_p = 500 cm^2/Vs in (5b) and (5c), which are roughly equal to the intrinsic values at room temperature for silicon [30,31]. For the quantum constants in (2a) and (2b), we used $m_n = 0.19 m_o$ (light electron mass in silicon) and $m_p = 0.49 m_o$ (heavy hole mass in silicon). These values result in a good match between DG simulations and experiment over a wide range of MOS device structures (see Section III.A). All simulations were for devices at room temperature (300K).

We implemented the above density-gradient model in a partial differential equation (PDE) solver called PROPHET [32]. This simulation tool provided several advantages over the traditional approach of coding a numerical simulator specifically for the DG model. It allowed us to specify the transport model in a script file at a high level to a general-purpose (but highly efficient) PDE solver. Also, PROPHET has the necessary facilities to solve the model in 1-D, 2-D, or 3-D for any specified device and test regime. Thus, the PDE-solver approach allows for the rapid investigation of a wide range of device structures, transport models and physical effects. Indeed, three significantly different versions of the DG model were investigated during this work. Using the conventional, model-specific approach, programming all of the major and minor model variations investigated would have taken many times as long. The classical QF model was also implemented in PROPHET, and was used for all classical model simulations.

III. Results and Discussion

A. Thin Oxide MOS Capacitors

The switching efficiency of a MOSFET is largely determined by its gate capacitance, which measures the ability of gate biases to control the carrier density and thus current flow below the gate oxide (see Figure 1). Thus, it is critical for simulations to accurately predict gate capacitance. The quantum repulsion of carriers from both gate and substrate oxide interfaces, as depicted in Figure 2, makes the oxide's effective electrical thickness to be typically 1-2 nm greater than its physical thickness. This effect is already quite noticeable in state-of-the-art commercial products, which have gate oxide thicknesses as low as 2 nm. The effect will quickly increase over the next decade, with gate oxide thicknesses predicted to shrink to 1 nm or less by 2012 [2].

To test this prediction, 1-D MOS capacitors with oxide thicknesses from 2 to 8 nm were simulated, and C-V curves were compared to those from classical simulations and from experimental measurements of the same structure [33]. As an example, the electron and hole densities for a 2.1 nm oxide device biased in inversion are shown in Figure 4. The resulting carrier density profiles were as sketched in Figure 2: near zero at the oxide interfaces, with the inversion or accumulation charge peak 0.5 to 1.5 nm beneath the Si-oxide interface, rather than exactly at the interface as in the classical model. Figure 3 compares C-V curves (capacitance versus gate bias V_G) for the 2.1 nm oxide MOS capacitor. As expected, the DG model reproduces measured data much more accurately than the classical (DD) model for this very thin oxide.

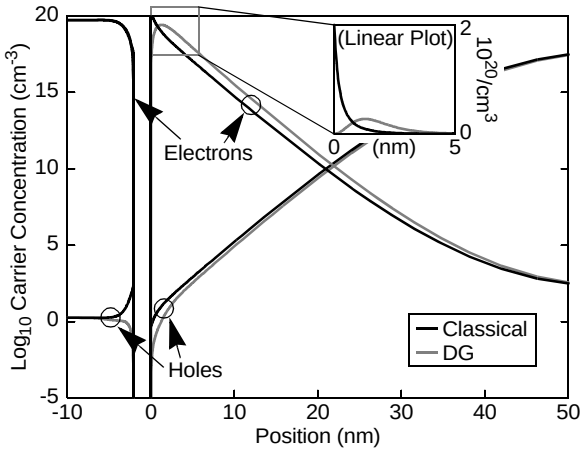


Figure 3: Classical and DG quantum-corrected carrier density profiles in a MOS capacitor operating in inversion. The inset details the electron profiles on a linear scale, showing the dramatic quantum correction at the oxide interface.

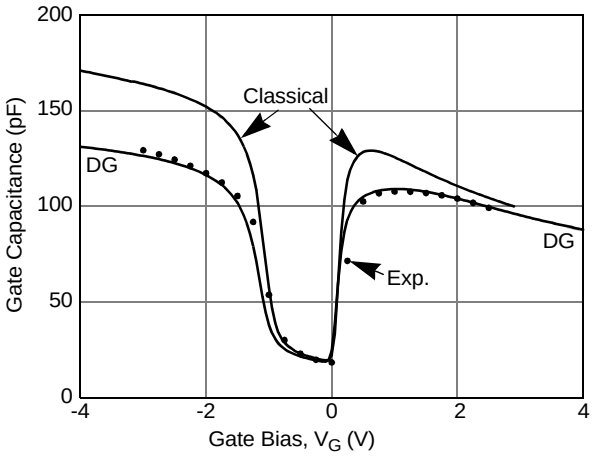


Figure 4: MOS capacitor low frequency C-V curve comparison for $(100\ \mu m)^2$ area, 2.1 nm thick oxide. The DG model reproduces measurements (dots; courtesy of H-P Labs) much more closely than the classical model.

To summarize the results of the C-V simulations over the range of oxide thickness considered, Figure 5 shows the fractional error in simulated capacitance (compared to measured data) versus oxide thickness for the classical and DG models. To simplify the plot, a single gate bias of V_G = -2V (accumulation) was chosen, since this condition is most critically affected by quantum effects and is least affected by other unknown parameters such as the poly doping level [34]. Here we see

that the DG model maintains accuracy at least down to 2 nm, while the accuracy of the classical model deteriorates rapidly for oxide thicknesses below 4 nm.

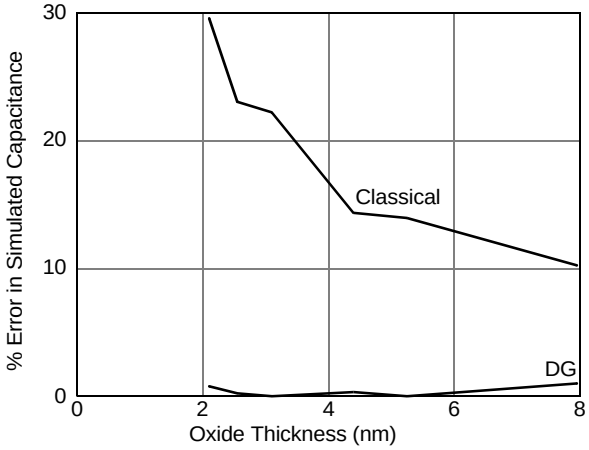


Figure 5: Percent error in simulated capacitance versus gate oxide thickness at V_G = -2V (accumulation). The DG model maintains accuracy at least down to 2.1 nm. The error of the classical model increases rapidly for oxide thicknesses below 4 nm.

B. Short Channel MOSFET

As discussed in Section I, 1-D simulations can provide only limited knowledge of the operation of real electronic devices. 2-D and 3-D simulations are often required for an accurate analysis of the operation of state-of-the-art (highly non-planar) devices. For example, the MOSFET operating current I_D, which flows horizontally from source to drain, is largely controlled by the vertical electric field from the gate. Thus, accurately modeling MOSFET operating current requires at least a 2-D analysis. Our previous work [34] was the first to show that 2-D simulations are quite feasible using the DG model. The fact that the DG model is also general (not structure-specific) allows it to work without modification or tuning for complex (e.g., non-planar) structures.

Figure 6 shows the simulated drain characteristic (drain current versus drain bias at a series of gate biases) for a very aggressively scaled MOSFET with a 30 nm gate length and 2 nm gate oxide thickness. This device approximates the state of the art in research labs [35]. Computed drain characteristics for both the quantum-corrected DG model and the classical quasi-Fermi model are shown. At each gate bias, the DG current is 20% to 70% below that predicted by the classical model. Thus, quantum effects result in a serious decrease in the current drive capability of the device. We point out, however, that even this aggressively scaled 30 nm MOSFET still performs like a switch. Figure 6 shows a negligible simulated drain current at logic-zero gate bias (0 V), and significant drain current at logic-one bias (1 V). Admittedly, the device is in need of extensive engineering to minimize short-channel effects.

One question which needs to be answered is whether the reduced DG current in Figure 6 is due to the reduction in channel charge, quantum transport effects along the channel, or both. Figure 7 compares the channel charge and drain current from the DG simulation (relative to the classical values) at full drain bias (1V) for the 30 nm MOSFET simulated in Figure 6. The close match between these curves over the full range of gate biases indicates that the DG current reduction is dominated by reduced channel charge, with only minor quantum transport effects. It makes sense that quantum effects are minor in the transport direction in the DG model, since this model

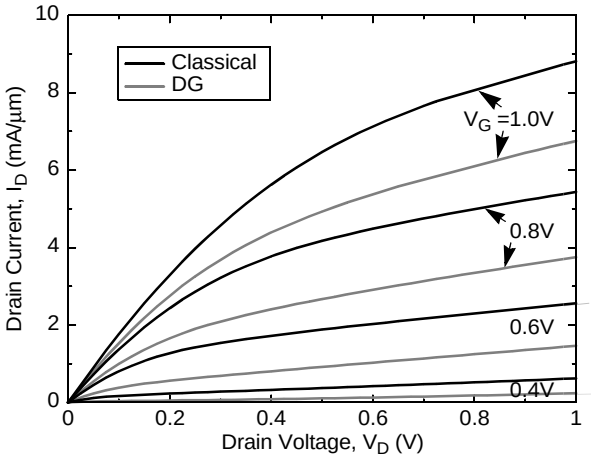


Figure 6: Simulated drain characteristics for 30 nm gate length, 2 nm gate oxide MOSFET. Results for classical and DG simulations are shown for 0.4V ≤ V_G ≤ 1V. The classical current is 20-70% larger.

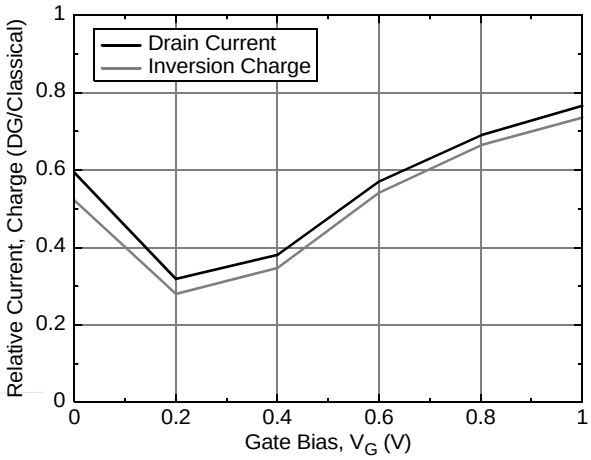


Figure 7: Relative drain current and channel charge (at mid channel with V_D = 1V) for the DG model (DG result divided by classical result). The result indicates that, in the normal operating range, the DG current reduction is dominated by reduced channel charge, with only minor quantum transport effects.

only significantly affects the potential and carrier profiles near abrupt heterojunctions and insulating interfaces. The relatively smooth potential in the transport direction results in small quantum potentials in this direction, and correspondingly small quantum effects on current. We note that other quantum models which include the effect of discrete quantum energy levels in the channel may predict more significant quantum effects due to carrier transport along the channel.

As final confirmation that the DG model and this tiny 30 nm MOSFET work correctly, Figures 8 and 9 compare classical and DG simulations of the subthreshold characteristic and DIBL (drain-induced barrier lowering versus channel length) for this device. In Figure 8, the subthreshold slope is 90 to 92 mV/decade of current for the classical model, and 105 to 110 mV/decade for the DG model. Thus, quantum confinement effects significantly degrade the subthreshold slope.

In Figure 9, DIBL was simulated by determining the threshold voltage versus channel length at maximum drain bias (V_D = 1V). To create longer channel MOSFETs, the 30 nm device was stretched at the center of the channel. The threshold voltage was computed as the gate bias at which the potential barrier to current flow between source and drain was the negative of the built-in potential in the substrate (0.494V in this case). In the classical model, this barrier is the classical potential at the gate oxide interface. In the DG model, the classical and quantum potentials combine to produce the barrier between source and drain, and the minimum barrier to carrier flow is slightly beneath the oxide surface. Figure 9 shows that the DG model shifts the threshold voltage by about 150 to 200 mV. Further, DIBL is somewhat worse (threshold voltage varies more rapidly) with the DG model. In both cases, variation with channel length is very severe near 30 nm for this device. Thus, although the 30nm MOSFET works, inevitable process variations would likely make this device structure unsuitable for ULSI.

IV. Computational Efficiency

The forgoing simulation results show that the DG model makes it *feasible* to include quantum effects accurately and generally in multi-dimensional electronic device simulation. In this section, we go further to show that the DG model is in fact quite *efficient* in accomplishing this. In partic-

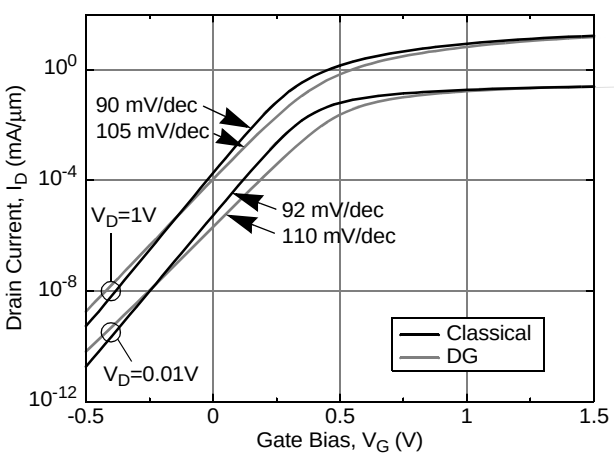


Figure 8: Subthreshold characteristic of 30-nm MOSFET. Quantum effects degrade the OFF state at low bias, the ON state at operating biases, and the saturation effect at high bias. The net effect is like a parallel conductance.

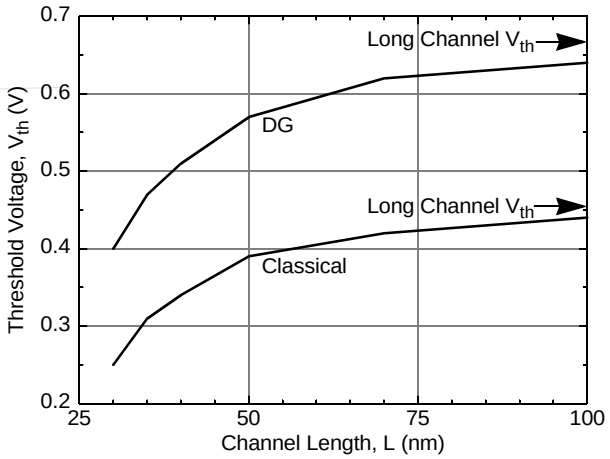


Figure 9: DIBL (drain-induced barrier lowering, or threshold voltage reduction) versus channel length. The channel of the 30-nm MOSFET was stretched to create longer channel MOSFETs. DIBL is somewhat more severe (stronger threshold variation with channel length) when quantum effects are included.

ular, we compare the computational cost of the classical and DG models for the simulations in Section III.

For the MOS capacitor simulations of Section III.A, we used a non-uniform 1-D grid with 320 to 450 points (depending on oxide thickness) for both classical and DG simulations. On an SGI O2 workstation, a typical C-V curve trace (81 bias points) took 40 seconds for the classical model, and 74 seconds for the DG model. For this case, computation time including quantum effects with the DG model is less than a factor of 2 larger than that for purely classical simulations.

For simulating MOSFET I-V curves in Section III.B, we used an identical 2-D grid with about 1750 points for both the classical and DG simulations. On the same workstation, a typical I-V curve trace (51 bias points) took 452 seconds with the classical model and 2383 seconds for the DG model. We note that the DG model in Section II proved to be unstable at times, in which case we used a slightly modified model with $\sqrt{n}$ and $\sqrt{p}$ as solution variables rather than n and p. An unresolved error in the Jacobian for the modified DG model results in linear convergence (rather than quadratic), and a correspondingly longer simulation time. Thus, computing an I-V curve with the modified DG model requires typically 5400 seconds.

These DG model computation times should be compared to the orders of magnitude increase in computation time for more rigorous quantum models, such as those based on the Schrödinger equation [36] or Green's functions [37]. Since the DG model is only moderately more computationally demanding than the associated classical models, it can even be feasibly solved in 3-D [38]. More importantly, the DG model leverages all of the accuracy and optimization tuning of the industry standard, classical drift-diffusion model. Thus, the DG model provides practical insight into quantum effects in ultra-small electronic devices.

V. Conclusions

In summary, we presented the density-gradient model as a computationally efficient means of including quantum effects in multi-dimensional electronic device simulation suitable for future

MOSFET technology. Computation time is typically 2 to 10 times that for the purely classical model, making it possible to run 2-D electronic device simulations routinely on a workstation. With the simulation of numerous 1-D and 2-D MOS devices, we also demonstrated the robustness of the DG model. In MOS capacitor simulations, classical model predictions rapidly diverge from measured results for oxide thicknesses below 4 nm, while the DG model maintains error below a few percent down to 2 nm.

In simulations of a 30 nm gate length MOSFET, quantum effects are predicted to reduce current drive by up to 70%. According to the DG model, this current reduction is almost entirely due to the reduced inversion charge resulting from vertical quantum confinement, while horizontal quantum transport effects along the channel are minimal. Finally, we showed that quantum effects degrade the subthreshold slope of this small MOSFET by 15-20 mV/decade, and increase DIBL. We conclude that the inclusion of quantum effects is essential for the accurate simulation of 30 nm scale MOSFETs, and that the density gradient model is an efficient way to accomplish that.

ACKNOWLEDGEMENTS

B. A. Biegel would like to express his appreciation for the technical assistance of Dr. T.R. Govindan. This work was supported in part by NASA Contract NAS2-14303.

REFERENCES

[1] *National Technology Roadmap for Semiconductors*, p. 13 (1997).
[2] *Ibid*., p. 46.
[3] E. Wang, *Second NASA Device Modeling Workshop*, Moffett Field, CA, Aug. 7-8, 1997.
[4] W. Shockley, *Bell Systems Technical Journal*, Vol. 28, p. 435 (1949). W. van Roosbroeck, *Bell Systems Technical Journal*, Vol. 29, p. 560 (1950).
[5] H. K. Gummel, *IEEE Transactions on Electronic Devices*, Vol. ED-11, p. 455 (1964).
[6] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, *Journal of Applied Physics*, Vol. 81, p. 7845 (1997).
[7] B. A. Biegel and J. D. Plummer, *Physical Review B*, Vol. 54, p. 8070 (1996), and references therein.
[8] R. K. Mains and G. I. Haddad, *Journal of Computational Physics*, Vol. 112, p. 149 (1994).
[9] W. Hänsch, T. Vogelsang, R. Kircher, and M. Orlowski, *Solid-State Electronics*, Vol. 32, p. 839 (1989).
[10] R. Rios, N. D. Arora, C.-L. Huang, N. Khalil, J. Faricelli, and L. Gruber, *1995 IEDM Technical Digest*, p. 937.
[11] M.J. van Dort, P. H. Woerlee, A. J. Walker, A. H. Juffermans, and H. Lifka, *IEEE Transactions on Electron Devices*, Vol. 39, p. 932 (1992).
[12] A. Spinelli, A. Benvenutti, and A. Pacelli, *1996 IEDM Technical Digest*, p. 399.
[13] J.B. Roldan, F. Gamiz, J.A. Lopez-Villaneuva, and P. Cartujo, *Journal of Vacuum Science and Technology B*, Vol. 16(3), p. 1538 (1998).
[14] J. P. Kreskovsky and H. L. Grubin, *VLSI Design*, Vol. 3(2), p. 179 (1995).
[15] D. K. Ferry and J.-R. Zhou, *Physical Review B*, Vol. 48(11), p. 7944 (1993).
[16] M. R. Pinto, *Comprehensive Semiconductor Device Simulation for VLSI*, p. 23, Ph.D. Thesis, Stanford University (Aug. 1990).
[17] M. G. Ancona and H. F. Tiersten, *Physical Review B*, Vol. 35, p. 7959 (1987).
[18] C. v. Weizsacker, *Zeitschrift fur Physik*, Vol. 96, p. 431 (1935).

[19] P. Hohenberg and W. Kohn, *Physical Review*, Vol. 136, p. 864 (1964).
[20] N.D. Mermin, *Physical Review*, Vol. 137, p. A1441 (1965).
[21] Another path of development which to a certain extent superseded the approach of density-gradient corrections for equilibrium systems is that of the Kohn-Sham equations that allow the direct inclusion of gradient effects to all orders.
[22] See, e.g., J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh and C. Fiolhais, *Physical Review B*, Vol. 46, p. 6671 (1992).
[23] F. Perrot, *Physical Review A*, Vol. 20, p. 586 (1979).
[24] Because fundamentally the DG term is a correction to the equation of state of the electron gas, one of us (MGA) prefers to describe the DG model as modifying the diffusion term of the DD model. Of course, mathematically this view is indistinguishable from that given in the text.
[25] F. Perrot, *Physical Review A*, Vol. 20, p. 586 (1979). M. G. Ancona and G. J. Iafrate, *Physical Review B*, Vol. 39, p. 9536 (1989).
[26] M. G. Ancona, Z. Yu, W.-C. Lee, R. W. Dutton, and P. V. Voorde, *Journal of Technology Computer Aided Design*, 1998 (http://tcad. stanford.edu/tcad-journal/archive).
[27] Actually, the quantum potentials would have to be infinite at the oxide interfaces to force the carrier densities to exactly zero there, so we instead set the densities at the interfaces to a small but non-zero value; 10-8/cm3 was used in this work. The particular small value chosen for carrier densithy in the oxides has a negligible effect on the simulation results.
[28] M. G. Ancona, Z. Yu, R. W. Dutton, P. J. Vande Vorde, M. Cao, and D. Vook, *IEEE Transactions on Electronic Devices*, Vol. ED-47, p. 2310 (2000).
[29] S. M. Sze, *Physics of Semiconductor Devices*, 2nd Ed., p. 84, Wiley, 1981.
[30] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, 2nd Ed., p. 55, Wiley, 1986.
[31] Exact current predictions are not being pursued here; only quantum effects on device behavior. More accurate mobility models are under development.
[32] PROPHET home page: http://www-tcad.stanford.edu/~conor/prophet/.
[33] Data courtesy of Hewlett-Packard Labs, Palo Alto. Capacitor areas were (100 μm)2. Poly doping in simulations was chosen between 3e19/cm3 and 5.5e19/cm3 to fit both DG and classical simulations to the slope of the measured C-V curve in inversion (VG > 1V).
[34] C. S. Rafferty, B. Biegel, Z. Yu, M. G. Ancona, J. Bude, R. W. Dutton, *Proceedings of SISPAD '98*, Leuven, Belgium, Sept. 2-4, 1998 p. 137.
[35] G. Timp et al., *1998 IEDM Technical Digest*, p. 615.
[36] A. Abramo, A. Cardin, L. Selemi, and E. Sangiorgi, "Two-Dimensional Quantum Mechanical Simulation of Charge Distribution in Silicon MOSFETs", *IEEE Transactions on Electron Devices*, Vol. 47(10), p. 1858 (2000).
[37] A. Svizhenko, M.P. Anantram, T.R. Govindan, B. Biegel, and R. Venugopal, *Journal of Applied Physics*, Vol. 91, p. 2343 (2002).
[38] A. Asenov, G. Slavcheva, A. R. Brown, J. H. Davies, and S. Saini, *1999 IEDM Technical Digest*, p.535.