Modeling of Quantum Effects for Ultrathin Oxide MOS Structures With an Effective Potential

Yiming Li, *Member, IEEE*, Ting-wei Tang, *Fellow, IEEE*, and Xinlin Wang

***Abstract*—In this paper, the effectiveness of the effective potential (EP) method for modeling quantum effects in ultrathin oxide MOS structures is investigated. The inversion-layer charge density and MOS capacitance in one-dimensional MOS structures are simulated with various substrate doping profiles and gate bias voltages. The effective mass is used as an adjusting parameter to compare results of the EP model with that of the Schrödinger–Poisson solution. The variation of this optimum parameter for various doping profiles at different gate voltages is investigated. The overestimated average inverse charge depth by the EP method is quantified and its reason explained. The EP model is a good practical simulation tool for modeling quantum effects but more work needs to be done to improve its accuracy near the interface.**

***Index Terms*—Effective potential, modeling and simulation, MOS devices, quantum effect, Schrödinger–Poisson.**

I. Introduction

As MOSFET devices are further scaled into the deep submicrometer regime, it has become necessary to include quantum mechanical effects when modeling device behavior. There have been two approaches to the modeling of these quantum effects: 1) employing full quantum mechanical transport model such as nonequilibrium Green's function [1] and 2) adding quantum corrections to the classical drift-diffusion (D-D) or hydrodynamic equation, such as the density gradient method [2].

Recently, a new effective potential (EP) approach has been advanced [3]–[6] which has the advantages of easy numerical implementation and almost guaranteed convergence. However, the accuracy of the EP model has never been systematically quantified against a more rigorous but time-consuming Schrödinger–Poisson (S-P) solution under various doping conditions. For this purpose, in this paper we have applied the EP method to the simulation of inversion-layer charge density (and therefore MOS capacitance as well) in one-dimensional (1-D) MOS structures with various substrate doping profiles and gate bias voltages. The effective mass is used as an adjusting parameter to compare results of the EP model with that of the S-P solution and the variation of this optimum parameter is investigated. The paper is organized as follows. Simulation procedures for the effective potential models are described in Section II. Section III presents the simulation results and discussion. Finally, Section IV draws conclusion.

Manuscript received June 8, 2002; revised October 11, 2002. This work was supported in part by the National Science Council of Taiwan under Contract NSC 91–2112-M-317–001, the National Science Foundation under Grant ECS-0120128, and the 2002 Research Fellowship Award of the Pen Wen-Yuan Foundation in Taiwan, R.O.C.

Y. Li is with the National Nano Device Laboratories, Hsinchu 300, Taiwan. He is also with the Microelectronics and Information Systems Research Center, National Chiao Tung University, Hsinchu 300, Taiwan, R.O.C. (e-mail: ymli@faculty.nctu.edu.tw).

T. Tang is with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA 01003 USA (e-mail: ttang@ecs.umass.edu).

X. Wang is with the Device and Process Modeling Group, IBM, Hopewell Junction, NY 12533 USA (e-mail: xinlinw@us.ibm.com).

Digital Object Identifier 10.1109/TNANO.2002.807386

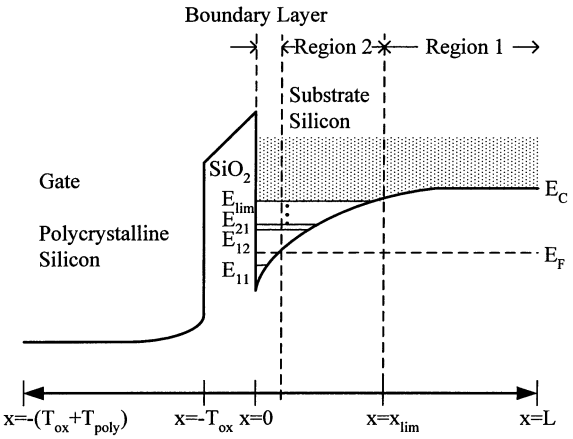

Fig. 1. A schematic band profile for the MOS structure.

II. Simulation Procedures

A polycrystalline-oxide-silicon MOS structure with a p-type silicon substrate as shown in Fig. 1 is simulated. An ideal oxide with a dielectric constant of 3.9 and a silicon substrate with a dielectric constant of 11.7 are assumed. Calculations are carried out at a temperature of 300 K using the D-D approximation and the Fermi–Dirac statistics. The oxide–silicon interface is chosen to be at $x = 0$. The oxide thickness is fixed at 1.6 nm. The silicon layer lies in the region $x > 0$ and the doping concentration of poly-silicon gate is 5×10^{19} cm^{-3}. Three different substrate doping profiles are considered, namely: 1) uniform doping of $N_A = 1 \times 10^{17}$ cm^{-3}; 2) low–high (retrograde) doping with $N_s = 1 \times 10^{17}$ cm^{-3} near the interface and abruptly rising to $N_{bulk} = 1 \times 10^{18}$ cm^{-3} at $x = 10$ nm; and 3) a Gaussian profile, $n(x) = N_{A0} + (D_I/\sqrt{2\pi\Delta R_P}) \exp(-(x - R_P)^2/2\Delta R_P^2)$ with $N_{A0} = 2 \times 10^{17}$ cm^{-3}, $D_I = 1 \times 10^{17}$ cm^{-3}, and $R_P = \Delta R_P = 10$ nm [7], [8].

First, we solve the S-P equations using 16 subbands approximation [7], [8]. A zero wave function boundary condition is forced at quantum system boundaries, i.e., at the bottom of the substrate and at the oxide interface. Therefore, penetration of the wave function into the oxide is neglected. This solution is used for comparison with the solution obtained by the EP model. The

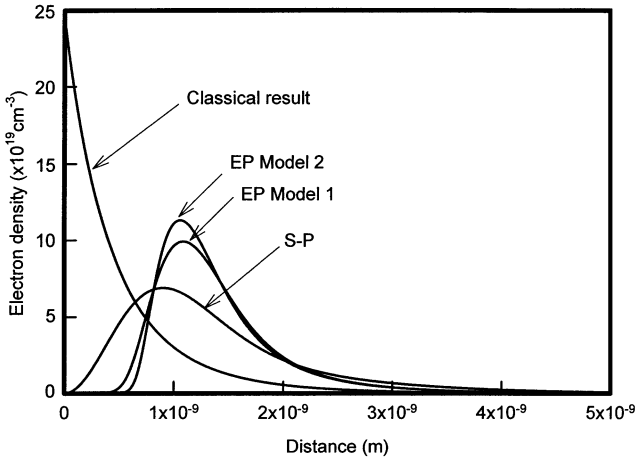

Fig. 2. Electron density with the S-P and two EP models, where the applied voltage is 1.5 V.

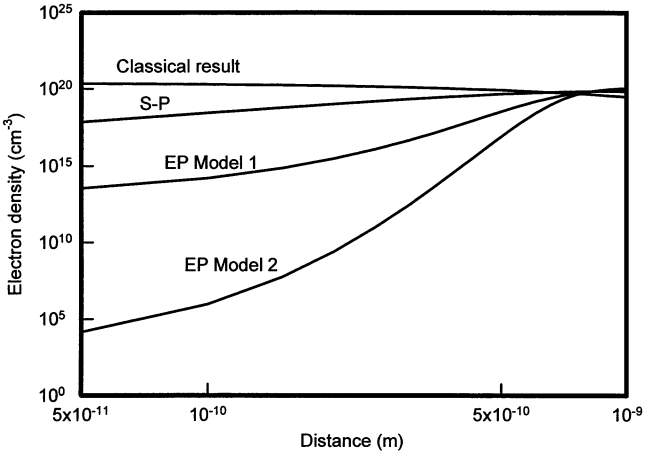

Fig. 3. Zoom-in plot of the electron density in log scale.

comparison criterion is based on the average inversion charge depth defined as

$$\langle x \rangle = \frac{\int_0^\infty x n(x) dx}{\int_0^\infty n(x) dx}. \tag{1}$$

The EP model is based on the following integral transformation from the classical potential (CP) $V(x)$ to the EP $V_{\text{eff}}(x)$ [6] as

$$V_{\text{eff}} = \frac{1}{\sqrt{2\pi}\alpha} \int_{-\infty}^{\infty} V(x+\xi) \exp\left(-\frac{\xi^2}{2a^2}\right) d\xi \tag{2}$$

where $a = \hbar/\sqrt{8m^* k_B T}$. The CP is initially calculated from the D-D equation coupled with Poisson's equation. After the EP is obtained, the electron concentration $n(x)$ is calculated using the Fermi–Dirac statistics. Poisson's equation is then solved again until a consistency is reached.

III. Simulation Results and Discussion

We present the calculated electron density $n(x)$ with the EP and S-P models using a (or equivalently the effective mass m^*) as an adjusting parameter to produce the $\langle x \rangle$ in both results. Two models for $V(x)$ are considered. Model 1 uses the electric potential obtained from Poisson's equation and it is a continuous function of position at the oxide–silicon interface. Model 2 uses $V(x) = E_c(x)/(-q)$ which has a discontinuity at the oxide–silicon interface. The latter seems to be a reasonable way of extending $V(x)$ to an infinite space rather than limiting it to the semi-infinite space on the silicon side only. Fig. 2 shows the comparison between results of the EP and S-P models together with the classical result. It is clearly seen that both EP models produce a peak electron density that is approximately twice as large as that of the S-P model, and the position of the peak density is further setback (1.0–1.2 nm) from the interface than that of the S-P model (0.8–1.0 nm). This has also been observed earlier by Watling *et al.* [9]. Fig. 3 presents a zoom-in plot of the electron density near $x = 0$ for these three models. Fig. 4 shows the corresponding effective potential calculated by the EP method and the classical potential by the S-P method, respectively.

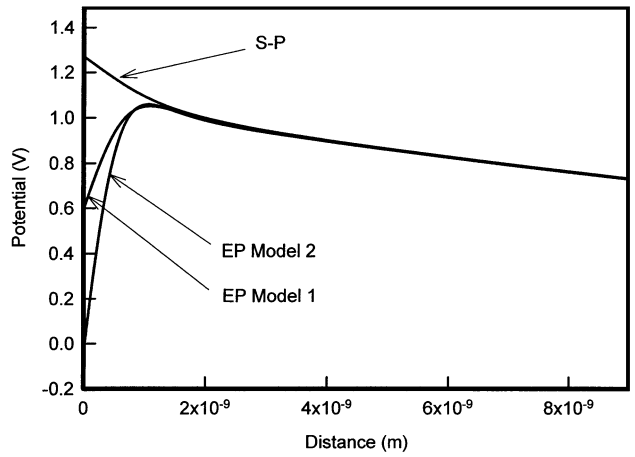

Fig. 4. Potentials for the S-P and EP models.

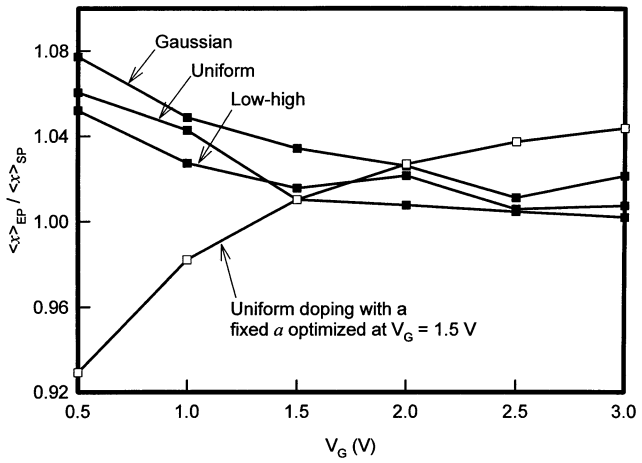

Fig. 5. Ratio of the $\langle x \rangle$ for the S-P and the EP model 1.

Fig. 5 shows the ratio of $\langle x \rangle_{\text{EP}}$ to $\langle x \rangle_{\text{SP}}$ with respect to the applied voltages for different doping profiles while keeping the same integrated inversion charge sheet density N_{inv} by adjusting the parameter a. The percentage derivation of $\langle x \rangle$ is large at low applied voltages ($\approx 5\%$) and becomes smaller at high voltages ($\approx 2\%$). Also shown in Fig. 5 (open squares) is the same ratio for the uniform doping if the parameter $a =$

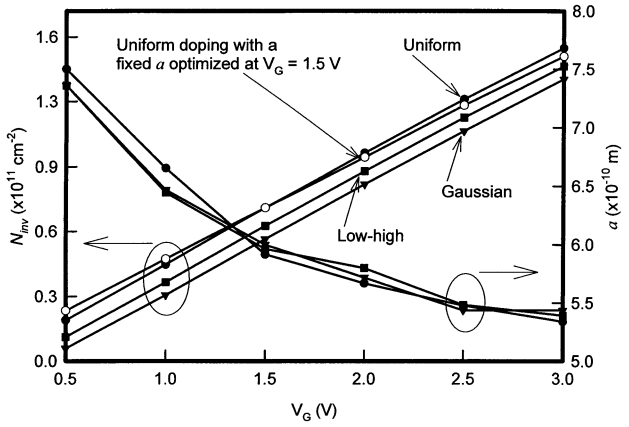

Fig. 6. Integration of the inversion layer charges and the a value for EP model 1.

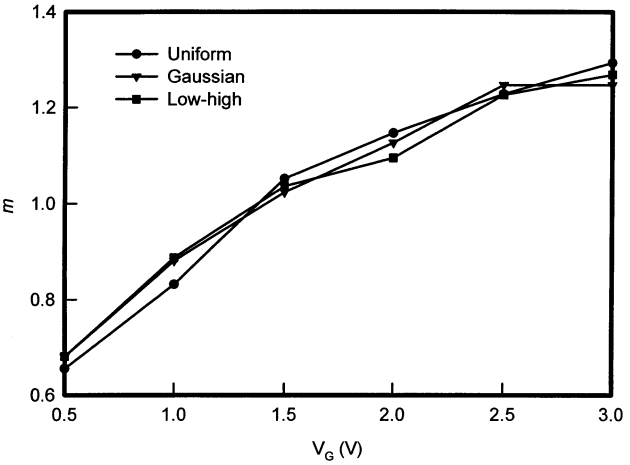

Fig. 7. Optimal m^* for the different doping profiles.

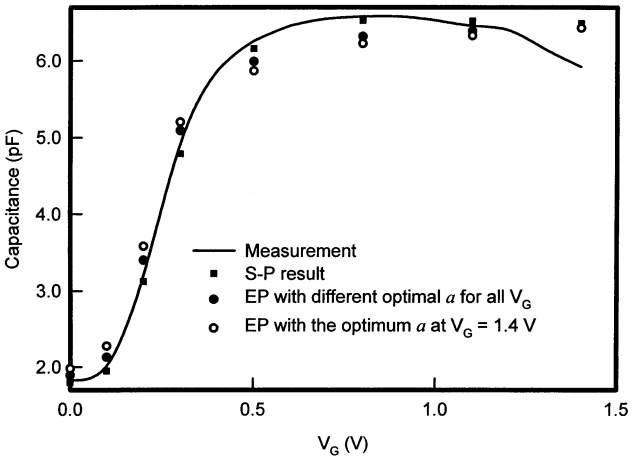

Fig. 8. Comparison of the measured *C–V* data with the S-P model and the EP model for all optimized a's and for using a fixed value a.

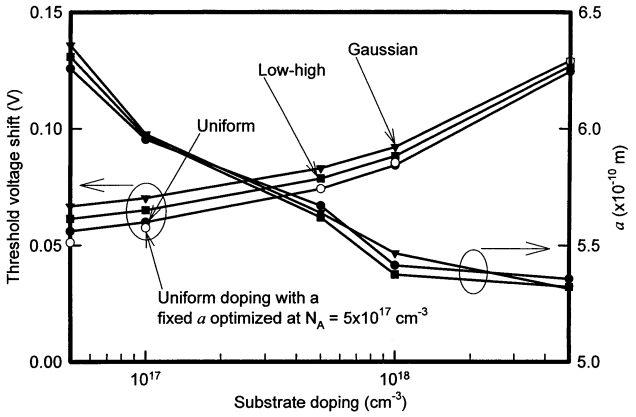

Fig. 9. Plot of the threshold voltage shift between the S-P solution and the classical result versus doping levels.

5.75×10^{-10} m optimized at $V_G = 1.5$ V is used for all other gate voltages. In this case, the range of the ratio is from 0.93 to 1.04.

Fig. 6 shows the optimum parameter a required to produce the same N_{inv} for both the EP and S-P models in Fig. 5. Also shown in Fig. 6 are these calculated N_{inv}'s for different doping profiles. The optimum parameter a to produce the same N_{inv}'s approximately varies from 7.5×10^{-10} m to 5.5×10^{-10} m for different V_G's. The open circles represent N_{inv} for the uniform doping if $a = 5.75 \times 10^{-10}$ m optimized at $V_G = 1.5$ V is used for all other gate voltages. In this case, the percentage errors in N_{inv} can be as large as 12% at $V_G = 0.5$ V.

Fig. 7 presents the adjusted effective mass m^* required to produce the same $\langle x \rangle$ as that of the S-P solution for various doping profiles. Note that, when the parameter a is optimized to produce the same $\langle x \rangle$ as that of the S-P solution, it automatically results in the same N_{inv} for both models. Using $V_G = 1.5$ V as a reference, the adjusted m^* varies as much as -40% at $V_G = 0.5$ V to $+25\%$ at $V_G = 3.0$ V.

In Fig. 8, a *C–V* curve measured at a frequency of 100 kHz of a 20×20 μm^2 N-MOS capacitor with $T_{\text{ox}} = 1.6$ nm is presented. Comparing the measured *C–V* data with the EP model, we have found that, through the calibration of a or m^*, the EP model can reproduce the experimental data fairly well. The maximum error of the result predicted by the EP model is only about 4.5% compared to the experimental data. Characteristically, the EP model underestimates the capacitance at high gate voltages. The main reason for this is that the total inversion charge sheet density predicted by the S-P and the EP models is about the same [10], although the average inversion charge depth is different. On the other hand, when the optimum parameter $a = 5.9 \times 10^{-10}$ m optimized at $V_G = 1.4$ V is used throughout, the maximum *C–V* error is about 10%.

We have also compared the difference in the threshold voltage using the classical, S-P, and EP models. The results are shown in Fig. 9 for a $T_{\text{ox}} = 1.6$ nm n-MOS structure which is similar to the result of Watling *et al.* [9]. The lines represent the threshold voltage shift $\Delta V_T = V_T(S-P) - V_T(\text{CL})$ and the symbols represent $\Delta V_T = V_T(\text{EP}) - V_T(\text{CL})$. Since the parameter a in the EP model is optimized to have the same integrated charge density as that of the S-P solution, the two models produce the same result and the symbols are on the top of the lines. However, if a fixed $a = 5.67 \times 10^{-10}$ m optimized for $N_A = 5 \times 10^{17}$ cm^{-3} is used for all other uniform doping concentrations, the voltage shift calculated by the EP model is then represented by open circles rather than by filled circles. As expected, the threshold voltage predicted by the quantum model is always higher than that of the classical model because of the

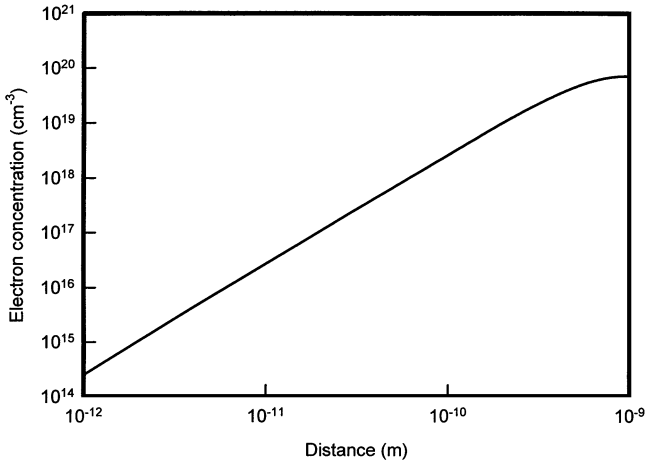

Fig. 10. Plot of the electron concentration of the S-P equations in the boundary layer, where $T_{\rm ox} = 1.6$ nm and $V_G = 1.5$ V.

quantization effect. Also shown in Fig. 9 is the optimized parameter a used in the EP model. The optimum parameter a to produce the same $N_{\rm inv}$'s approximately varies from 6.3×10^{-10} m to 5.2×10^{-10} m for different doping profiles and levels.

The idea of proposing the EP model for nanoscale device simulation originates in part from the consideration of the finite-size effect of charge carriers [11], [12]. When the electric potential produced by a point charge (represented by a delta function) is replaced by a finite-size wave packet (represented by a Gaussian function), the CP is smoothed out and it results in an EP as given by (2). The integral form of $V_{\rm eff}$ as given in (2) suggests that $V_{\rm eff}$ is a convolution of $V(x)$ with a Gaussian function in an infinite space. In the original paper by Feynman and Kleinert [13], the EP was used to study bounded carriers in a symmetric anharmonic oscillator and a double-well potential. Since the relationship between $V_{\rm eff}(x)$ and $V(x)$ in (2) is linear, both potentials are supposed to satisfy the same boundary conditions. However, the quantum boundary conditions imposed on $V_{\rm eff}(x)$ may be different from the classical boundary conditions imposed on $V(x)$. For example, along the longitudinal channel direction in a MOSFET, both $V_{\rm eff}$ and V satisfy the same classical boundary conditions at the source and drain ends. In this case, the EP model smoothes out and lowers the potential barrier at the source–channel junction resulting in a slight increase in the drain current [3]. However, in the transverse direction underneath the MOS capacitor under the strong inversion condition, the discontinuity in $E_c(x)$ at the interface is abrupt and quite large (≈ 3.1 eV) which causes the confinement of charge carriers. In this case, the tunneling or penetration of the electron wavefunction at the interface may require more than simple smoothing of $V(x)$ by the Gaussian transformation. In addition, because the Gaussian transformation in (2) considers primarily the electron finite-size effect in an infinite space, when it is applied to a semi-infinite space problem, such as one at the oxide–silicon interface, it may not be the most appropriate form to describe the effect of the finite extent of the wave packet.

In our earlier study [7], [8], we have found that, for a nontunneling boundary condition at the oxide–silicon interface, a plot of $\log n(x)$ versus $\log x$ shows a clear linear slope of 2 (see Fig. 10). This implies that $n(x) \propto x^2$ as $x \to 0$ for an infinite high oxide barrier. In the EP model, whether it is Model 1 or Model 2, the plot of $\log n(x)$ versus x is almost linear (see Fig. 3). This means that $n(x)$ is an exponential function of position because $V_{\rm eff}(x)$ is almost a linear function of x near the neighborhood of $x = 0$ as seen in Fig. 4. The electron density near the interface is reduced in the EP model because $V_{\rm eff}$ is lower than the classical potential V near the surface. This is a result of either the semi-infinite space integration of $V(x)$ in Model 1, or using $V(x) = E_C(x)/(-q)$ to include the discontinuity at the interface in Model 2. In essence, the Gaussian distribution may be a good recipe for bounded carriers in a symmetric potential well but it may not be the most appropriate form for bounded carriers in a semi-infinite triangular well.

IV. Conclusion

In this paper, we have investigated the quantum effects occurring at the oxide–silicon interface under the inversion condition with the EP method for the ultrathin oxide MOS structure. The range of the optimum parameter a is determined for various doping profiles at different gate voltages. Using a as an adjusting parameter, the EP model can produce a correct amount of the charge sheet density but its prediction of the average inversion charge depth is always overestimated. This introduces a small error in the *C–V* measurement curves. The optimum parameter a is also somewhat sensitive to the doping concentrations and doping profiles. Nevertheless, the EP method is a convenient, if not the most accurate, way of producing a first-order result for the quantization effects in semiconductors.

Acknowledgment

The authors express their appreciation to the referee for an exceptional indepth reading of the manuscript.

References

[1] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *J. Appl. Phys.*, vol. 81, no. 12, pp. 7845–7869, June 1997.

[2] M. G. Ancona, "Equations of state for silicon inversion layers," *IEEE Trans. Electron Devices*, vol. 47, pp. 1449–1456, July 2000.

[3] D. K. Ferry, "The onset of quantization in ultra-submicron semiconductor devices," *Superlatt. Microstruct.*, vol. 27, no. 2/3, pp. 61–66, Feb. 2000.

[4] D. K. Ferry, R. Akis, and D. Vasileska, "Quantum effects in MOSFETs: Use of an effective potential in 3D Monte Carlo simulation of ultra-short channel devices ," *IEDM Tech. Dig.*, pp. 287–290, 2000.

[5] S. M. Ramey, R. Akis, D. Vasileska, and D. K. Ferry, "Modeling of quantum effects in ultrasmall SOI MOSFET's with effective potentials," in *Abst. of Silicon Nanoelectronics Workshop*, 2001, pp. 50–51.

[6] L. Shifren, R. Akis, and D. K. Ferry, "Correspondence between quantum and classical motion: Comparing Bohmian mechanics with a smoothed effective potential approach," *Physics Letters A*, vol. 274, pp. 75–83, Sep. 2000.

[7] X. Wang, "Quantum Correction to the Charge Density Distribution in Inversion Layers," M.S. thesis, Univ. Massachusetts, Amherst, 2001.

[8] X. Wang and T.-W. Tang, "Comparison of three quantum correction models for the charge density in MOS inversion layers," J. Comput. Electron., submitted for publication.

[9] J. R. Watling, A. R. Brown, A. Asenov, and D. K. Ferry, "Quantum corrections in 3-D drift diffusion simulations of decanano MOSFET's using an effective potential," in *Simulation of Semiconductor Processes and Devices 2001*, D. Tsoukalas and C. Tsamis, Eds. Vienna, Austria: Springer-Verlag, 2001, pp. 82–85.

[10] R. Akis, L. Shifren, D. K. Ferry, and D. Vasileska, "The effective potential and its use in simulation," *Phys. Stat. Sol. (b)*, vol. 226, no. 1, pp. 1–8, 2001.
[11] D. K. Ferry and H. L. Grubin, "Electrons in semiconductors: How big are they?," in *Extended Abstracts of 1998 6th Int. Workshop on Computational Electronics*, 1998, pp. 84–87.
[12] D. K. Ferry, "Simulation at the start of the new millennium: Crossing the quantum-classical threshold," *VLSI Design*, vol. 13, no. 1-4, pp. 155–161, 2001.
[13] R. P. Feynman and H. Kleinert, "Effective classical partition functions," *Phys. Rev. A*, vol. 34, pp. 5080–5084, 1986.

Yiming Li (M'02) received the B.S. degrees in applied mathematics and electronics engineering, the M.S. degree in applied mathematics, and the Ph.D. degree in electronics all from National Chiao Tung University, Hsinchu, Taiwan, in 1996, 1998, and 2001, respectively.

In 2001, he joined the National Nano Device Laboratories (NDL) as an Associate Researcher and the Microelectronics and Information Systems Research Center as as Assistant Professor, National Chiao Tung University (NCTU), where he has been engaged in the research and development of modeling and simulation of nanodevices and VLSI circuits. He was a Visiting Assistant Professor at the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, in the fall of 2002. He is in charge of the nanodevice TCAD Laboratory and the Parallel and Scientific Computing Laboratory at NDL and NCTU. His current research areas include semiconductor nanostructures physics, modeling, and device simulation, VLSI circuit simulation, computational biology, and advanced numerical techniques and scientific computation. He has authored or coauthored over 90 researcher papers in international books, journals, and conferences.

Dr. Li is a member of Phi Tau Phi. He received a Research Fellowship Award from the Pan Wen-Yuan Foundation, Taiwan in 2002. Dr. Li is a member of the Institute of Electronic, Information and communication Engineers, the American Physical Society, the Society for Industrial and Applied Mathematics, and WSEAS.

Ting-wei Tang (SM'70-F'99) received the B.S. degree in electrical engineering from the National Taiwan University, Taiwan in 1957, and the M.Sc. and Ph.D degrees in engineering from Brown University, Providence, RI, in 1961 and 1964, respectively.

From 1964 to 1968, he was an Assistant Professor of Aerospace Engineering at the University of Connecticut, Storrs. In 1968, he joined the Electrical and Computer Engineering Department, University of Massachusetts, Amherst, as an Associate Professor and was promoted to Professor in 1974. He was a Visiting Professor at Oxford University, Oxford, U.K., in 1975 and a Visiting Samsung Professor of Electronics Engineering at Seoul National University, Seoul, Korea in the fall of 1990. During 1999–2000, he was a Research Chair at the National Center for High-Performance Computing, Hsinchu, Taiwan. His current research interest is in semiconductor device physics and numerical modeling of semiconductor devices. He teaches a graduate-level course in which students gain hands-on experience in developing an elementary device simulation software. He has published more than 100 journal articles and conference papers.

Dr. Tang served on the IEEE Admission and Advancement Committee (1974–1977), the Technical Program Committee of IEEE BCTM (1995–1998), and the Scientific Advisory Committee of ISSDT (1994–1995). He received an Outstanding Senior Faculty Award and an Outstanding Teaching Award from the College of Engineering, University of Massachusetts in 1989 and 1991, respectively, and Chancellor's Medal from the University of Massachusetts in 2000. He was a keynote speaker at the 1993 Symposium on Semiconductors Modeling and Simulation in Taipei, Taiwan, the 1995 Fourth International Seminar on Simulation of Devices and Technologies at Kruger National Park, South Africa, and the 1999 Semiconductor Technology CAD Workshop and Exhibition in Hsinchu, Taiwan.

Xinlin Wang received the B.S. degree in electronics engineering from Tsinghua University, China, in 1998 and the M.S. degree in electrical engineering from the University of Massachusetts, Amherst, in 2001.

Currently she is working in the TCAD group in the Semiconductor Research and Development Center of IBM, Hopewell Junction, NY. Her research interests focus on device physics, modeling, and simulation of nanoscale transistors.