PAPER *Special Issue on the IEEE International Conference on SISPAD '02*

On Density-Gradient Modeling of Tunneling through Insulators

Timm HÖHR[†a)], **Andreas SCHENK**[†], **Andreas WETTSTEIN**[††], ***and*** **Wolfgang FICHTNER**[†], ***Nonmembers***

SUMMARY The density gradient (DG) model is tested for its ability to describe tunneling currents through thin insulating barriers. Simulations of single barriers (MOS diodes, MOSFETs) and double barriers (RTDs) show the limitations of the DG model. For comparison, direct tunneling currents are calculated with the Schrödinger-Bardeen method and used as benchmark. The negative differential resistance (NDR) observed in simulating tunneling currents with the DG model turns out to be an artifact related to large density differences in the semiconductor regions. Such spurious NDR occurs both for single and double barriers and vanishes, if all semiconductor regions are equally doped.

key words: *device simulation, density-gradient model, direct tunneling, gate leakage, resonant tunneling*

1. Introduction

Quantum effects in modern deep-submicron devices are of growing interest. A prominent unwanted quantum effect in MOSFETs is direct tunneling through the thin gate dielectric, which increases the off-state power consumption. Another important effect is quantum depletion, the carrier density decay towards barriers. In MOSFETs this leads to a shift of the inversion charge maximum away from the oxide interface which in turn causes a shift of the threshold voltage, a lowering of the gate capacitance and an apparent increase of the oxide thickness.

There are several methods that can be combined with conventional drift diffusion simulators to include these effects: Quantum depletion effects can be accurately modeled by solving the one-dimensional Schrödinger equation along the confinement direction self-consistently coupled with the device equations [1].

Established methods for modeling direct tunneling [2] are the calculation of a transmission coefficient [3] and the use of Bardeen's transfer Hamiltonian [4], [5] with either quasi-classical Wentzel-Kramers-Brillouin (WKB) wave functions or self-consistently obtained numerical solutions of the 1D-Schrödinger equation [1].

An interesting, computational efficient alternative for including quantum effects into conventional device simulators is the density gradient (DG) model. This model introduces a quantum correction term containing higher derivatives of the density or the electrostatic potential into the usual drift-diffusion or hydrodynamic device equations [6]–[9]. The DG model is known to describe quantum depletion effects very well [1]. It also has been applied to one-dimensional insulator tunneling [10] (using two carrier populations according to tunneling direction) and source-to-drain tunneling in ultrashort channel MOSFETs [11].

The aim of this paper is to demonstrate the degree to which the DG transport model is capable of reproducing direct tunneling currents through insulating barriers. As a reference we use simulations solving the one-dimensional Schrödinger equation combined with Bardeen's transfer Hamiltonian method for calculating the tunneling current [1], [4]. All devices studied here are silicon based with single or double SiO_2 barriers.

The DG model and modifications for non-equilibrium are described in Sect. 2. Simulations of tunneling characteristics with these models and the reference method are presented in Sect. 3. The findings are discussed in Sect. 4.

2. Model

The density gradient (DG) model (or 'quantum drift diffusion,' QDD) [1], [7], [9], [12], [13] can be viewed as a modification of the usual drift diffusion (DD) model. A 'quantum potential' Λ is introduced into the classical formulas of the electron density n and the current density $\vec{J}_n$ (we restrict the considerations below to electrons, for holes corresponding expressions exist):

$$n = N_\mathrm{c} \exp\left[\beta\left(E_{\mathrm{F},n} - E_\mathrm{c} - \Lambda\right)\right] \tag{1}$$

$$\begin{aligned}\vec{J}_n &= -\mu\, kT\, \nabla n - \mu\, n\, \nabla(E_\mathrm{c} + \Phi_m + \Lambda) \\ &= -\mu\, n\, \nabla E_{\mathrm{F},n},\end{aligned} \tag{2}$$

with $\beta = 1/kT$, the electron quasi-Fermi energy $E_{\mathrm{F},n}$, the conduction band edge E_c and a mass driving term $\Phi_m = -kT \log N_\mathrm{c}$ from DOS discontinuities. In the formulation presented here, Λ is the solution of the partial differential equation:

Manuscript received September 4, 2002.

[†]The authors are with Institut für Integrierte Systeme, ETH Zürich, Gloriastrasse 35, CH-8092 Zürich, Switzerland.

[††]The author is with ISE Integrated Systems Engineering AG, Balgriststrasse 102, CH-8008 Zürich, Switzerland.

a) E-mail: hoehr@iis.ee.ethz.ch

$$\Lambda = -\frac{\gamma\hbar^2}{12\,m}\left[\nabla^2 \log n + \frac{1}{2}\left(\nabla \log n\right)^2\right] \tag{3}$$

$$= -\frac{\gamma\hbar^2\beta}{12\,m}\left[\nabla^2\left(\xi E_{\mathrm{F},n} - \overline{\Phi}\right) + \frac{\beta}{2}\left(\nabla\left(\xi E_{\mathrm{F},n} - \overline{\Phi}\right)\right)^2\right], \tag{4}$$

where $\xi = 1$, γ is a fit factor and $\overline{\Phi} = E_{\mathrm{c}} + \Phi_m + \Lambda$. Together with Poisson and continuity equation the formulas (1,2,4) form a system that has to be solved self-consistently.

The density gradient model is derived as a moment expansion of the Wigner-Boltzmann [14] or a corresponding quantum Liouville equation [9] and a situation close to thermal equilibrium is assumed. The derivation in [1] calculates an equilibrium solution for the density matrix as a first order perturbation in E_c. Additionally it is assumed that E_c varies only slowly on the scale of the thermal de-Broglie wavelength. The result is a formulation given by (1,2,3) with $\gamma = 1$ where instead of the density n actually the *classical* equilibrium density $n_{\mathrm{cl}} \propto \exp(-\beta E_c)$ appears in (3). When this is replaced by the *quantum mechanical* density n we obtain (3) with Λ appearing also on the right hand side. This replacement generates a smooth $\overline{\Phi}$ even at band edge jumps $\Delta E_c \gg k_B T$, which violates the assumptions made above, but corresponds to the situation encountered at the Si-SiO_2 interface. This step still lacks a satisfactory justification. Nevertheless the DG model is able to describe equilibrium densities in MOS channels very well when compared to a more accurate 1D-Schrödinger solver [1].

'Tunneling' in the DG formulation originates from the reduction of the barrier† by the presence of Λ in the current equation (2). The carriers only have to surmount the residual barrier $\overline{\Phi}$. The tunneling current is not separated from the drift-diffusion current and hence is determined by the mobility μ_{ox} in the oxide, which we use as a fitting parameter. For each material a constant mobility was used.

In order to allow for transport modeling, a space dependent quasi-Fermi energy is introduced in (4). This expression can be generalized by setting $\xi \neq 1$ [15]. As the derivation of Eq. (3) is valid close to equilibrium only, the proper value for ξ is not known from theory. For tunneling, $E_{\mathrm{F},n}$ varies significantly over the barrier, and the value for ξ matters. Therefore, we examine the cases $\xi_{\mathrm{ox}} = 1$ and $\xi_{\mathrm{ox}} = 0$ in oxide regions. In semiconductor regions, $E_{\mathrm{F},n}$ varies little and we use $\xi = 1$ throughout.

3. Simulations and Results

Single and double barrier devices were studied with the simulation tool DESSIS$_{-\mathrm{ISE}}$. The implementation of the density gradient model is described in [15]. As a reference, the one-dimensional Schrödinger equation is solved self-consistently with the Poisson equation in a region covering oxide and part of the substrate next to it. The direct tunneling current is then calculated by the Bardeen method a posteriori using the numerically computed wave functions and assuming plane waves in the polysilicon gate [1].

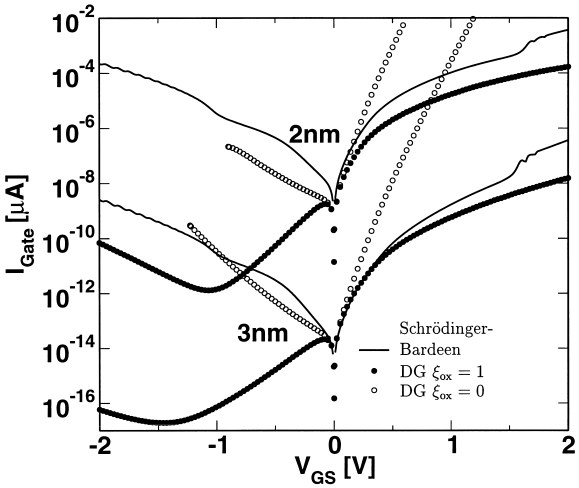

Fig. 1 Gate 'tunneling' currents in n-channel MOSFETs. Density gradient results (symbols) are compared to Schrödinger-Bardeen (lines).

3.1 N-Channel MOSFET

Gate tunneling characteristics (gate current I_{Gate} versus gate voltage V_{GS}) were produced for symmetric n-channel MOSFETs†† with two oxide thicknesses (Fig. 1). Source, drain and back contact were kept at zero potential.

For $\xi_{\mathrm{ox}} = 1$ and small positive bias ($V_{\mathrm{GS}} \lesssim 0.5\,\mathrm{V}$) one obtains DG curves close to Schrödinger-Bardeen (SB) results by using an oxide mobility $\mu_{\mathrm{ox}} = 0.05\,\mathrm{cm}^2/\mathrm{Vs}$. However, for $V_{\mathrm{GS}} < 0$ there is a strong discrepancy, the most peculiar feature being a current peak very close to 0 V and a minimum enclosing a region $-1\,\mathrm{V} \lesssim V_{\mathrm{GS}} \lesssim 0\,\mathrm{V}$ where negative differential resistance (NDR) occurs.

Using $\xi_{\mathrm{ox}} = 0$ yields monotonously rising currents, which are, however, too high for positive and too low for negative bias. Hence, fitting μ_{ox} does not improve the situation. Only in the very vicinity of $V_{\mathrm{GS}} = 0$ (equilibrium) DG 'tunneling' currents match the reference curves given by the SB method.

For the 2 nm device additional SB simulations have been carried out including a self-consistent current calculation. Apart from a worse convergence behavior no difference was found for the I_{Gate}-V_{GS} characteristics.

†The insulator is treated as a semiconductor with a wide band gap and insulator parameters.

††For all simulated devices the term MOS actually implies a highly n-doped ($10^{20}\,\mathrm{cm}^{-3}$) polysilicon region instead of metal.

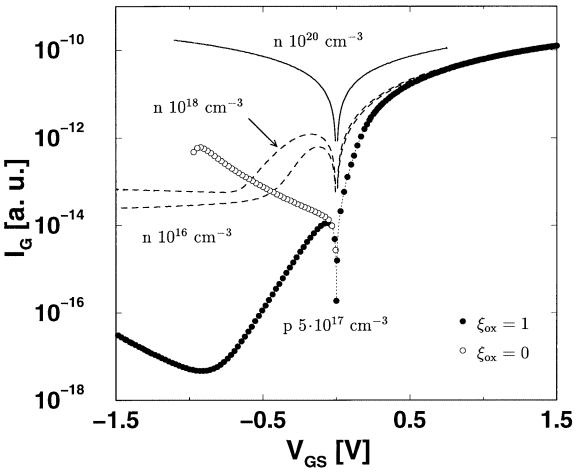

Fig. 2 DG 'tunneling' currents for MOS diodes (structure n^+Polysilicon-Oxide-Si) with different Si dopings. All curves are shown for $\xi_{ox} = 1$ unless indicated otherwise.

The current is too small to add a significant contribution to the substrate space charge.

3.2 MOS-Diode

For the simpler structure of a one-dimensional MOS-diode with 2 nm oxide thickness, similar DG current characteristics are obtained as for the MOSFET (symbols in Fig. 2). The label V_{GS} now applies to the voltage at the n^+-polysilicon 'gate' contact with respect to the substrate. Here, having no source and drain contact, the carrier supply is limited by thermal generation. For a better comparability, the lifetimes of SRH generation/recombination in the substrate were set to extremely small values. Exploring the case $\xi_{ox} = 1$ for different substrate doping (Fig. 2) we find that the NDR behavior becomes weaker, if the p-type substrate is changed to n-type and finally vanishes for symmetric doping, as expected for a symmetric device structure.

The electron density n and the residual barrier $\overline{\Phi}$ are shown for $\xi_{ox} = 0$ and $\xi_{ox} = 1$ in Figs. 3(a) and (b), respectively. The inset in Fig. 3(b) compares the conduction band edge E_c with the effective band edge $\overline{\Phi}$ for the case of thermodynamic equilibrium. The barrier is largely reduced. In equilibrium the two cases for ξ_{ox} are equivalent, but with ceasing inversion they exhibit different profiles in the oxide as well as in the substrate region next to it. Most striking is the discontinuity of the density at the oxide-silicon interface for $\xi_{ox} = 0$.

3.3 Resonant Tunneling Diode

NDR is an effect known to occur in resonant tunneling devices (RTDs). The results with single barrier MOS-structures motivated the investigation of the DG model applied to silicon RTDs with two SiO_2 barriers enclosing a quantum well of varying thickness. The structure

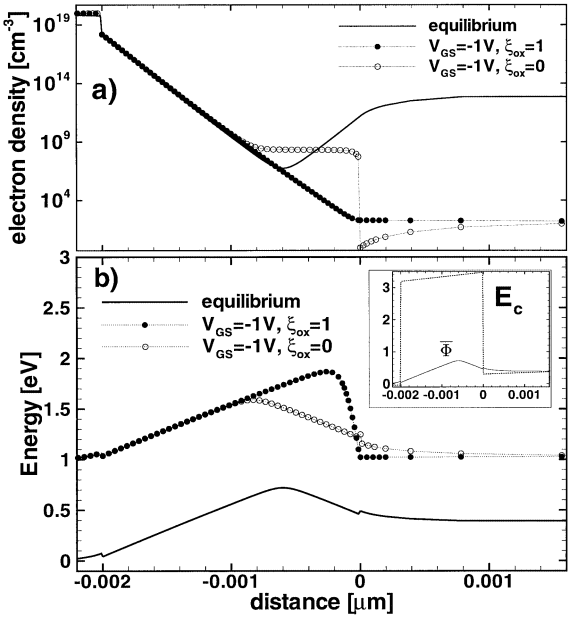

Fig. 3 Electron density n (a) and effective band edge $\overline{\Phi}$ (b) along a MOS diode at different gate voltages for $\xi_{ox} = 1$ (filled symbols) and $\xi_{ox} = 0$ (open symbols). The structure (from left to right) is: n^+-Polysilicon-Oxide-Si. The oxide-semiconductor interfaces are located at 0 and −2 nm, respectively. The inset in b) shows the equilibrium effective barrier $\overline{\Phi}$ compared to the conduction band edge E_c. The small steps in $\overline{\Phi}$ at the interfaces are due to the DOS discontinuities Φ_m that are not included in this graph.

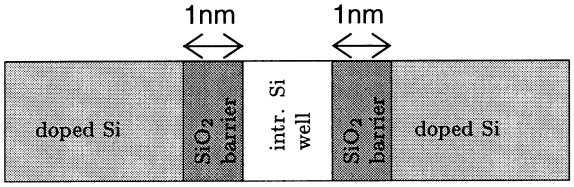

Fig. 4 Structure of a RTD as used in the simulations. The well consists of an intrinsic silicon region sandwiched between two SiO_2 barriers of 1 nm width.

is shown in Fig. 4. The well is intrinsic and the outer regions are highly n-doped (10^{20}cm^{-3}). The barriers are 1 nm wide. For all following results $\xi_{ox} = 1$ was used.

Current characteristics obtained from DG simulations are shown in Fig. 5 (dashed lines). In addition, a curve for a single oxide barrier between intrinsic and n-doped silicon is included (circles in Fig. 5) which seems to be approached, if the well length is increased in the double barrier device. Furthermore, the NDR-like feature vanishes, if the outer regions and also the well are equally n-doped (solid line in Fig. 5).

The occurrence of the DG current peak and a corresponding NDR is related to the dimension of the intrinsic well region. It is present if the well extends over 5 nm or more. For a narrow well, measuring only 1 nm, this effect does not appear (thin dashed line in

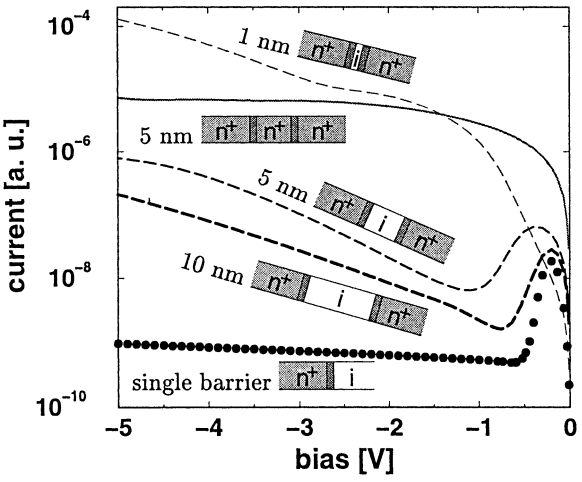

Fig. 5 Currents for RTDs with different well lengths calculated with the DG model ($\xi_{ox} = 1$). The small pictures illustrate the device structure. There are three kinds: The RTDs have an intrinsic well with different lengths (dashed lines, white middle regions). One RTD has a n-doping of $10^{20}\mathrm{cm}^{-3}$ also in the well (solid line, shaded middle region). The third structure is a single barrier MOS-diode with an intrinsic substrate (•).

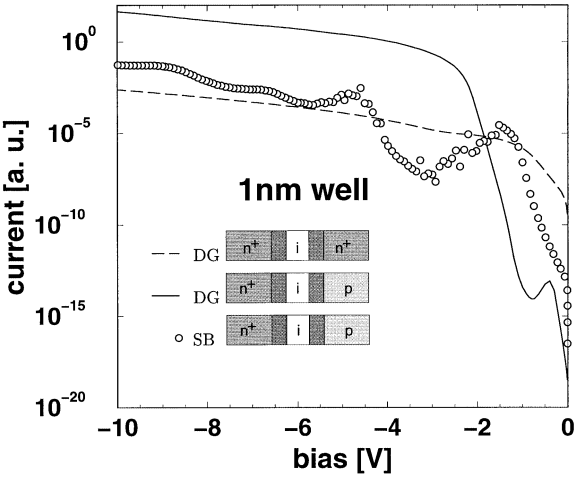

Fig. 6 DG current characteristics for an RTD with asymmetrical (p-i-n^+) doping (solid line) compared to Schrödinger-Bardeen (symbols). A symmetrically doped device (n^+-i-n^+) is also shown (dashed lines). Well and barriers are both 1 nm wide.

Fig. 5). For this small well length NDR reappears only by switching to p-doping in one of the outer regions, i.e. when the difference in density across the *whole* device is increased, which is shown in Fig. 6 (solid line).

In Fig. 6 characteristics of SB and DG simulations are compared for a RTD having a p-doped and a n^+-doped electrode and an intrinsic well with a length of 1 nm. The SB characteristics exhibits two main resonance peaks (open circles in Fig. 6) that are clearly different in number, location and peak value from the single current peak that is obtained for the same device with the DG model (solid line in Fig. 6).

For the corresponding symmetrically n^+-doped device with an intrinsic well, the peak is absent (dashed line in Fig. 6). As seen before, this is related to the small well dimension. This curve, like for a WKB approximation, shows no resonance but seems to depict an average of the SB characteristics.

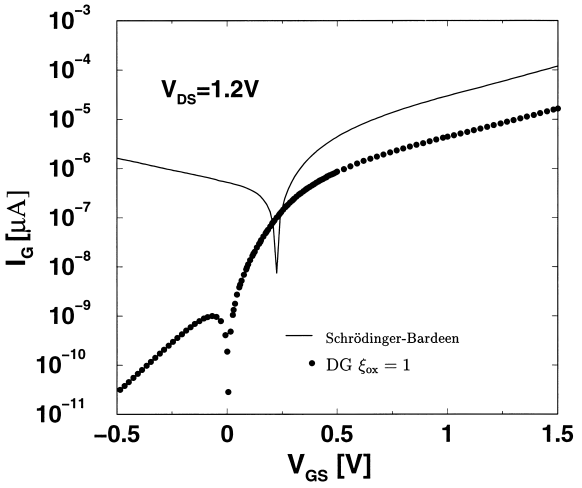

Fig. 7 N-MOSFET gate leakage current as a function of gate voltage V_{GS} for a drain voltage of $V_{DS} = 1.2\,\mathrm{V}$. The oxide thickness is 2 nm. Density gradient results using $\xi_{ox} = 1$ (symbols) are compared to Schrödinger-Bardeen (line).

3.4 N-MOSFET Off-State Leakage

The question to which extent the DG model may describe the tunneling contribution to off-state leakage is of great interest for industrial application. Therefore, simulations were performed at a fixed source-drain voltage ($V_{DS} = 1.2\,\mathrm{V}$) using either the Density Gradient model with $\xi_{ox} = 1$, or the Schrödinger-Bardeen approach (Fig. 7). For gate voltages larger than 0.5 V the DG characteristics qualitatively follows the SB reference at a current level lower by almost one decade. Here a better fit could be obtained by adjusting the oxide mobility. In contrast, the currents differ dramatically in the low and negative bias region. This behaviour is fully consistent with the earlier observation that the characteristics are best reproduced if the electron density is high on both sides of the barrier and that channel depletion is accompanied by the occurrence of spurious NDR (see results for $V_{DS} = 0$ and MOS-diode).

It is also striking that the reference model shows a clear positive shift of the point of zero gate current compared to the DG result. At this point the gate-to-drain tunneling, which is already present in the off-state, is counterbalanced by an opposite component from source to gate that rises with increasing gate potential. The DG model is unable to reproduce the correct position of this point.

4. Conclusion

The DG model has been used to simulate electron tunneling across oxide barriers in silicon MOSFETs, MOS-diodes and RTDs. The modified model ($\xi_{ox} = 0$) produces discontinuous carrier densities, if tunneling occurs from high to low density regions. Non-monotonous current-voltage curves are observed for standard ($\xi_{ox} = 1$) DG simulations of single barrier as well as double barrier structures. In a MOSFET this behaviour prevents a qualitative reproduction of the off-state tunneling current.

The negative differential resistance vanishes, if both sides of a barrier are symmetrically n-doped or bias conditions are such that high electron densities exist on both sides (inversion). Only in this case reasonable IV-curves turn out, which show similarities to those obtained with the WKB approximation. Thus, the presence of spurious NDR is related to large density differences across the heterostructure.

Particularly for RTDs a NDR-like feature in the DG simulation disappears, if all semiconductor regions are equally doped. If there were a resonance peak, however, symmetric doping would only slightly change the peak position due to a shift of the bottom of the well. Furthermore, the resonances resulting from a Schrödinger-Bardeen simulation of a RTD completely differ in location from that of the spurious DG current peak. Therefore, the latter is not related to resonant tunneling. The similarities between single and double barriers also indicate that these features are not caused by quantum interference. This is also not expected, since the DG model does not contain any phase information.

Acknowledgment

This work was in part supported by the Kommission für Technologie und Innovation under contract 4082.2.

References

[1] A. Wettstein, Quantum Effects in MOS Devices, Hartung-Gorre Verlag, Konstanz, 2000.

[2] R. Clerk, A. Spinelli, G. Ghibaudo, and G. Pananakakis, "Theory of direct tunneling current in metal-oxide-semiconductor structures," J. Appl. Phys., vol.91, pp.1400–1409, 2002.

[3] A. Schenk and G. Heiser, "Modeling and simulation of tunneling through ultra-thin gate dielectrics," J. Appl. Phys., vol.81, pp.7900–7908, 1997.

[4] J. Bardeen, "Tunnelling from a many-particle point of view," Phys. Rev., vol.6, pp.57–59, 1961.

[5] W.A. Harrison, "Tunnelling from an independent-particle point of view," Phys. Rev., vol.6, pp.85–89, 1961.

[6] M.G. Ancona and H.F. Tiersten, "Macroscopic physics of the silicon inversion layer," Phys. Rev. B, vol.35, pp.7959–7965, 1987.

[7] M.G. Ancona and G.J. Iafrate, "Quantum correction to the equation of state of an electron gas in a semiconductor," Phys. Rev. B, vol.39, pp.9536–9540, 1989.

[8] C.L. Gardner, "The quantum hydrodynamic model for semiconductor devices," SIAM J. Appl. Math., vol.54, pp.409–427, 1994.

[9] C.L. Gardner and C.A. Ringhofer, "The smooth quantum potential for the hydrodynamic model," Phys. Rev. E, vol.53, pp.157–167, 1996.

[10] M.G. Ancona, "Equation of state for silicon inversion layers," IEEE Trans. Electron Devices, vol.47, no.7, pp.1449–1456, 2000.

[11] A. Schenk and A. Wettstein, "2d analysis of source-to-drain tunneling in decananometer mosfets with the density gradient model," Proc. 5th Int. Conf. Modeling and Simulation of Microsystems, ed. M. Laudon and B. Romanowicz, pp.552–555, San Juan, 2002.

[12] J.R. Zhou and D.K. Ferry, "Ballistic phenomena in gaas mesfets: Modelling with quantum moment equations," Semicond. Sci. Technol., vol.7, pp.B546–B548, 1992.

[13] ISE Integrated Syst. Eng. AG, "Dessis 7.0 reference manual," 2001.

[14] E. Wigner, "On the quantum correction for thermodynamic equilibrium," Phys. Rev., vol.40, pp.749–759, 1932.

[15] A. Wettstein, O. Penzin, and E. Lyumkis, "Integration of the density gradient model into a general purpose device simulator," VLSI Design, vol.15, no.4, pp.751–759, 2002.

Timm Höhr received the diploma in physics from the Universität Konstanz in 2000. In 2001 he started as Ph.D. student at the Integrated Systems Laboratory at ETH Zurich.

Andreas Schenk was born in 1957. He received the Dipl. Phys. degree and the Ph.D. from Humboldt University in Berlin (HUB) in 1981 and 1987, respectively. From 1987 till 1991 he was working on various aspects of the physics and simulation of optoelectronic devices. In 1991 he joined the Integrated Systems Laboratory of ETH working as a senior research/teaching assistant, where he qualified to give lectures at university in 1997 for "Physics and Modeling of Micro-electronic Devices." His main activities are in the physics-based modeling for advanced simulation of sub-micron silicon devices and their application in the TCAD software released by ISE AG Zurich. His interests include generation-recombination, mobility, contacts, heterojunctions, many-body effects, hot carrier degradation, noise, single electron transistor modeling at device level, and quantum effects in silicon ultra-small devices. He authored and co-authored 2 books and 80 papers. Andreas Schenk is a member of the German Physical Society (DPG).

Andreas 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 he is with ISE AG, Zürich, working as a software developer for device simulation.

Wolfgang Fichtner received the Dipl.Ing. degree in physics and the Ph.D. degree in electrical engineering from the Technical University of Vienna, Austria, in 1974 and 1978, respectively. From 1975 to 1978, he was an Assistant Professor in the Department of Electrical Engineering, Technical University of Vienna. From 1979 through 1985, he worked at AT&T Bell Laboratories, Murray Hill, NJ. Since 1985 he is Professor and Head of the Integrated Systems Laboratory at the Swiss Federal Institute of Technology (ETH). In 1993, he founded ISE Integrated Systems Engineering AG, a company in the field of technology CAD. Wolfgang Fichtner is a Fellow of the IEEE, a member of the Swiss National Academy of Engineering, and a corresponding member of the Austrian Academy of Sciences. In 2000, he received the IEEE Grove Award.