On the Interpretation of Local Negative Mobilities in Nanoscale Semiconductor Devices

Simon C. Brugger and Andreas Schenk

Abstract—In a previous theoretical work, the concept of mobility has been unequivocally extended to inhomogeneous non-equilibrium systems. This generalization naturally suggests a new one-particle Monte Carlo method to solve the Boltzmann equation, which can self-consistently take into account generation–recombination processes, as well as quantum corrections. This new scheme has been successfully applied to different kinds of MOSFETs. The results of the simulations clearly show that, surprisingly, the mobility in the channel can become negative. In this brief, we present a detailed analysis of this phenomenon and show that negative mobilities are directly related to regions, where quasi-ballistic transport takes place.

Index Terms—Boltzmann equation (BE), inverse scattering operator (ISO), negative mobility, quasi-ballistic transport.

I. INTRODUCTION

IN AN effort to properly model submicrometer silicon-on-insulator (SOI) MOSFETs, where the combination of hot-electron effects with generation–recombination processes can play a dominant role [3], a new one-particle Monte Carlo (OPMC) method has been developed [2]. This iterative method, which we called current-based OPMC method, uses transport parameters (TP) extracted from MC solutions of a Boltzmann equation (BE) in the considered device, as described in [2]. The extracted TPs (mobilities and diffusivities) are, in contrast to other methods (e.g., [4], [5]), always well and unequivocally defined, not only in one dimension but also for 2-D and 3-D systems. During the early stage of validation of our new method, we simulated small bulk and SOI MOSFETs. The results of those simulations clearly show that, surprisingly, the local mobility in transport direction can take negative values. After repeatedly verifying this observation,[1] we had to find an explanation for it. This brief presents a detailed analysis of the origin and the meaning of negative mobilities in nanoscale MOSFETs. In Section I, we recall the general definition of the mobility and how it is related to other more popular defini-

[1]Early results published in [1] have been successfully reproduced by the group of Jungemann [6] by directly solving the BE.

tions of this TP. Then, the physical meaning of the mobility is discussed. In Section II, we will present two examples of devices, where negative mobilities appear. Section III gives a detailed analysis of the negative mobility found in a simple NIN structure. Section IV generalizes the findings of Section III and shows how this phenomenon is directly related to regions where strong quasi-ballistic transport takes place. Section V will conclude the discussion.

II. THEORY

When trying to derive an equation for the electron current density from the BE

$$\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{r}} f - \frac{q}{\hbar}\mathbf{E} \cdot \nabla_{\mathbf{k}} f = Sf \tag{1}$$

where f is the distribution function, S the scattering operator, q the absolute value of the electron charge, $\mathbf{v}$ the velocity, and $\mathbf{E}$ the electric field, the first difficulty comes from the fact that the distribution function does not appear on its own in the equation (for a complete discussion, see [1]). Usually, to avoid this problem, the so-called relaxation-time approximation (RTA) is performed. The RTA consists in replacing the complicated scattering term Sf by the more simple expression

$$-\frac{f - f_{\text{eq}}}{\tau} \tag{2}$$

where τ is the relaxation time, and f_{eq} is the equilibrium distribution function. Then, multiplying (2) and the left-hand side of (1) by $q\mathbf{v}(\mathbf{k})\tau$ and integrating over all momenta $\mathbf{k}$, an approximate equation for the current density $\mathbf{J}_n$ is obtained

$$\mathbf{J}_n(\mathbf{r}) \simeq \int_K \mathbf{v}(\mathbf{k})\tau(\mathbf{r},\mathbf{k})\partial_t f(\mathbf{r},\mathbf{k})d^3k$$

$$+ q\int_K \mathbf{v}(\mathbf{k})\tau(\mathbf{r},\mathbf{k})\mathbf{v}(\mathbf{k}) \cdot \nabla_{\mathbf{r}} f(\mathbf{r},\mathbf{k})d^3k$$

$$- \frac{q^2}{\hbar}\int_K \mathbf{v}(\mathbf{k})\tau(\mathbf{r},\mathbf{k})\nabla_{\mathbf{k}} f(\mathbf{r},\mathbf{k}) \cdot \mathbf{E}(\mathbf{r})d^3k. \tag{3}$$

The mobility is traditionally defined as being the term in the front of the electric field [7] divided by the electron charge and the electron density

$$\mu_{ij} := -\frac{q}{n\hbar}\int_K \mathbf{v}_i(\mathbf{k})\tau(\mathbf{r},\mathbf{k})\nabla_{\mathbf{k}_j} f(\mathbf{r},\mathbf{k})d^3k. \tag{4}$$

Equation (4) is, however, not exact, because it was derived using the approximation (2) whose validity is difficult to ensure, particularly far from thermodynamic equilibrium. In [8], an exact expression for the current density is formally derived by inverting the scattering operator S. The inverse scattering operator (ISO) S^{-1} has the nice property [8]

$$S^{-1}Sf = f - f_{\text{eq}}. \tag{5}$$

By letting the ISO act on both side of (1) and then contracting with $-q\mathbf{v}(\mathbf{k})$, an exact expression for the current density $\mathbf{J}_n$ can be derived

$$\mathbf{J}_n(\mathbf{r}) = -q \int_K S_{\mathbf{v}}^{-1}(\mathbf{r}, \mathbf{k}) \partial_t f(\mathbf{r}, \mathbf{k}) d^3k$$

$$- q \int_K S_{\mathbf{v}}^{-1}(\mathbf{r}, \mathbf{k}) \mathbf{v}(\mathbf{k}) \cdot \nabla_{\mathbf{r}} f(\mathbf{r}, \mathbf{k}) d^3k$$

$$+ \frac{q^2}{\hbar} \int_K S_{\mathbf{v}}^{-1}(\mathbf{r}, \mathbf{k}) \nabla_{\mathbf{k}} f(\mathbf{r}, \mathbf{k}) \cdot \mathbf{E}(\mathbf{r}) d^3k \tag{6}$$

where $S_{\mathbf{v}}^{-1}(\mathbf{r}, \mathbf{k})$ is the velocity moment of the ISO, i.e., the vector with ith component

$$S_{\mathbf{v}_i}^{-1}(\mathbf{r}, \mathbf{k}) := \int_K \mathbf{v}_i(\mathbf{k}') S^{-1}(\mathbf{r}, \mathbf{k}', \mathbf{k}) d^3k'. \tag{7}$$

According to the traditional definition of mobility exposed above, one finds the exact expression

$$\mu_{ij} := \frac{q}{n\hbar} \int_K S_{\mathbf{v}_i}^{-1}(\vec{k}, \vec{r}) \nabla_{\mathbf{k}_j} f(\vec{k}, \vec{r}) d^3k. \tag{8}$$

The mobilities presented in this brief are all computed using the exact expression (8). In the case of parabolic band structures and/or purely elastic scattering, (8) reduces to (4), and the RTA is exact (see Appendix).

One could argue that (8) is quite abstract, that it allows no physical understanding of the mobility, and that a more physical definition is needed. To counter such arguments, we would like to recall that, even in simple systems, the physical meaning of the mobility is less clear than one would usually allege. To do so, we begin with the only case where the mobility has a clear physical meaning.

A. Bulk Systems

By applying a homogeneous electric field $\mathbf{E}$ to a homogeneous bulk system, the relation between the mean particle (here electrons) velocity $\mathbf{v}$ and electric field is given by

$$\mu\mathbf{E} = -\mathbf{v}. \tag{9}$$

Equation (9) leads to the most physical interpretation of the mobility, which is found in any textbook, i.e., "the parameter

relating carrier drift velocity and electric field" (see, e.g., [9]). The electric field is often called driving force (DF). In the bulk case, the DF has a profound physical meaning because it enters the equation for the momentum variation in time

$$\frac{d\mathbf{k}}{dt} = -\frac{q}{\hbar}\mathbf{E} \tag{10}$$

where q is the absolute value of the electron charge.

Considering simple inhomogeneous systems close to the thermodynamic equilibrium already leads to a mobility concept with less physical meaning.

B. Inhomogeneous Systems Close to Thermodynamic Equilibrium

Considering a 1-D inhomogeneous electron system close to thermodynamic equilibrium, one can use the Einstein relation and the variable transformation

$$n = n_i \exp\left(\frac{q(\phi - \psi)}{k_B T}\right) \tag{11}$$

where ϕ is the electrostatic potential, ψ is the quasi-Fermi potential, T is the thermodynamic temperature, and k_B is the Boltzmann constant, to rewrite the equation for the mean particle velocity $\mathbf{v}$ as (see, e.g., [1, pp. 110–111])

$$\mu(-\nabla_r\psi) = -\mathbf{v}. \tag{12}$$

Comparing (12) with (9), one can see that the only difference is that the gradient of the electrostatic potential has been replaced by the gradient of the quasi-Fermi potential. Because of this analogy, the gradient of the quasi-Fermi potential is also called DF. This DF, however, has less physical meaning than its conterpart in (9), because no equivalent to (10) exists. This is a first example, where, to keep the physical meaning of the mobility, a nonphysical DF is introduced.

C. General Case

In the general case,[2] the equation for the mean velocity is

$$\mu\mathbf{E} + \frac{\left(\nabla_r^T(\mathbf{D}^T n)\right)^T}{n} = -\mathbf{v} \tag{13}$$

where μ and $\mathbf{D}$ are the mobility and the diffusivity tensor, as defined in [8]. The physical meaning of the mobility in (13) is

---

[2]Here, "general case" means that the scattering operator is a piecewise constant function of the position as is always implicitly assumed in the Monte Carlo simulator. In the case of a continuous dependence on the position, an additional **diffusion** term appears in (13) [10]. This additional term does, however, change nothing to our argumentation.

far from being clear. In order to keep the usual interpretation, one can define a generalized DF $\mathbf{F}$ as

$$\mathbf{F} := \mathbf{E} + \mu^{-1} \frac{\left(\nabla_r^T(\mathbf{D}^T n)\right)^T}{n}. \qquad (14)$$

This is, however, not satisfying, because it only shifted the nonphysical meaning to the DF.

Thus, in the general case, one should go back to the general definition (8), i.e., the mean value of the projection of the $\mathbf{k}$ gradient of the distribution function on $S_{\mathbf{v}}^{-1}$.

III. NEGATIVE MOBILITIES

A. Band and Scattering Models

The results presented in this section were obtained using our in-house MC simulator SimnIC. The first four conduction bands were taken into account, computed with the empirical pseudopotential method (see [11], [12]). For the scattering operator S, the following model has been used.

1) Phonon scattering (as described in [13]).
   a) One acoustic (6.3 meV) intravalley, intraband, and interband.
   b) Two acoustic (12.1 and 18.5 meV) g-type intervalley, intraband, and interband.
   c) One optical (62.1 meV) g-type intervalley, intraband, and interband.
   d) Two acoustic (19.0 and 47.4 meV) f-type intervalley, intraband, and interband.
   e) One optical (59.0 meV) f-type intervalley, intraband, and interband.
2) Impurity scattering (elastic intravalley, intraband scattering as described in [14]).
3) Impact ionization (intervalley, intraband, and interband scattering, as described in [15]).

Using this model, the ith component of the velocity moment of the ISO takes the form [1]

$$S_{\mathbf{v}_i}^{-1}(\mathbf{r}, \mathbf{k}, v, b) = \frac{\check{\mathbf{v}}_i\left(\epsilon_b(\mathbf{k}), v, b\right) - \mathbf{v}_i(\mathbf{k}, b)}{W_{\text{tot}}\left(\mathbf{r}, \epsilon_b(\mathbf{k}), v, b\right)}$$
$$+ \check{S}_{\mathbf{v}_i}^{-1}\left(\mathbf{r}, \epsilon_b(\mathbf{k}), v, b\right) \qquad (15)$$

where v is the valley index, b is the band index, $\epsilon_b(\mathbf{k})$ is the energy at position $\mathbf{k}$ in the bth band, and $\check{g}(\epsilon_b(\mathbf{k}), v, b)$ is the mean value of the function $g(\mathbf{k})$ on an isoenergy surface at energy $\epsilon_b(\mathbf{k})$ in the vth valley and in the bth band.

B. First Example

The first example is a simple NIN structure composed of two N-regions of 0.1 μm with a doping concentration of 5×10^{18} cm^{-3} and an intrinsic region with a length of 0.14 μm. At a bias of 2 V, the mobility becomes negative 25 nm after the NI transition on a length of 8 nm, as can be seen in Fig. 1. Fig. 1 also shows the profile of the electric field. In the zone of negative mobility, a strong electric field prevails. Because of
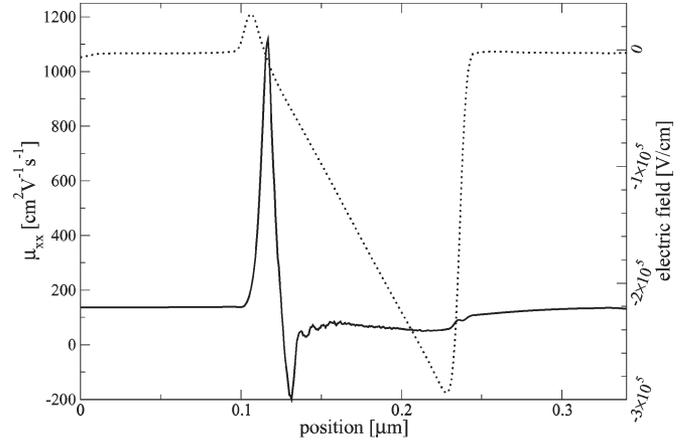

Fig. 1. Profile of the electric field (dotted line) and of the mobility in the transport direction.
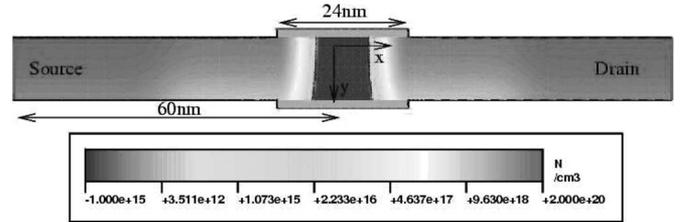

Fig. 2. Nano double-gate MOSFET and its doping.

this strong field, one can directly suspect that this could be a quasi-ballistic effect.

At this bias point, it has been observed [1, pp. 219–223] that, astonishingly, the two functions $\check{S}_{\mathbf{v}_i}^{-1}$ and $\check{\mathbf{v}}_i$ play a negligible role. Thus, in that case, (4), together with the relaxation time, $\tau(\mathbf{r}, \mathbf{k}) := 1/W_{\text{tot}}(\mathbf{r}, \mathbf{k})$ are a very good approximation of (8). Equation (4) will, therefore, be used instead of (8) to simplify the discussion in the next section.

C. Second Example

The second example is the double-gate MOSFET shown in Fig. 2. The effective gate length is 24 nm, and the body thickness 12 nm. In [2], we showed that, for a gate voltage of 1.1 V and a drain voltage higher than 0.6 V, the mobility (and even the effective mobility) in parts of the channel becomes negative. Again, looking at the profile of the electric field, one observes the same features as in the previous example.

To get a better physical understanding of this phenomenon, we depicted in Fig. 3 the projection of the 3-D distribution function in the $k_x k_y$ plane of the $\mathbf{k}$ space (the darker the zone, the higher the probability). In the device, the electric field is pointing in the $\langle 110 \rangle$ direction. Fig. 3 clearly shows that the distribution function in the different valleys is shifted toward the $\langle 110 \rangle$ direction, and that, it is strongly anisotropic. Fig. 4 shows the contribution to the negative mobility from the different $\mathbf{k}$ points (the darker, the more negative). Comparing Figs. 3 and 4, one finds that the negative maxima of the mobility are directly correlated with the maxima of the distribution function.
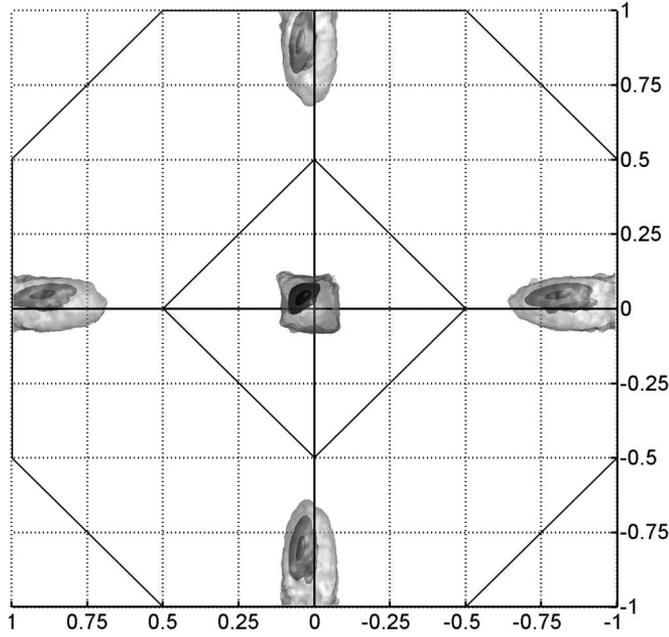
Fig. 3.   Isosurfaces of the distribution function (k_x x-axis and k_y y-axis).
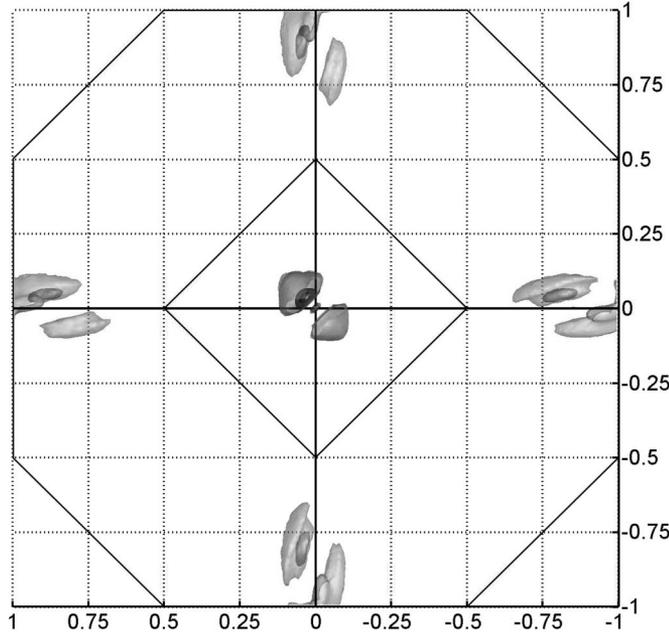
Fig. 4.   Isosurfaces of the mobility (k_x x-axis and k_y y-axis).
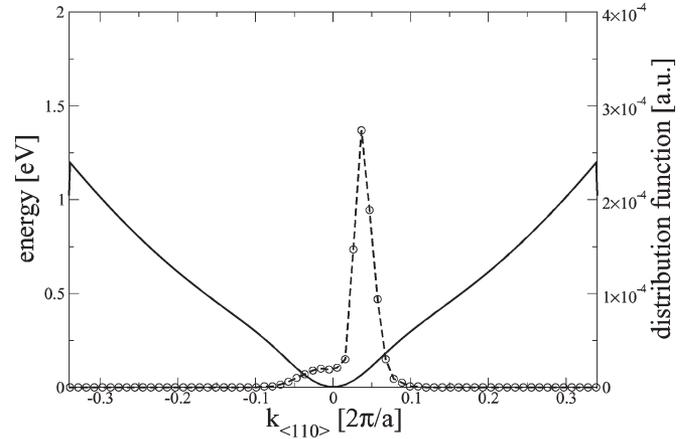

Fig. 5.   One-dimensional cuts of the distribution function (dashed line) and the band structure.
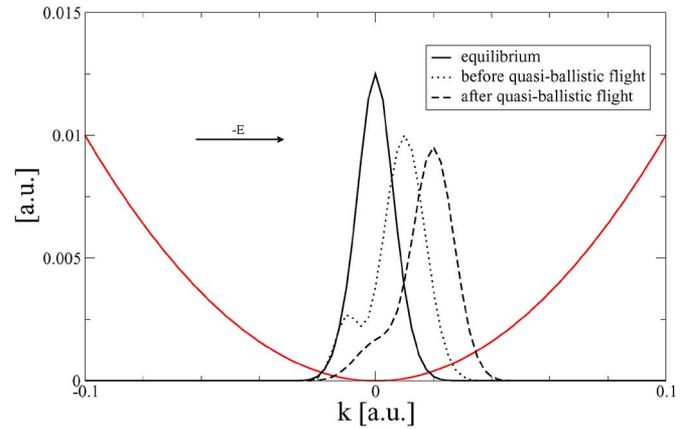
Fig. 6.   Illustration of the distribution function in three regions of importance and the band structure.

IV. INTERPRETATION

The 1-D cut in the distribution function, along the $\langle 110 \rangle$ direction at $k_z = 0.83 \cdot 2\pi/a$, shown in Fig. 5, gives a first insight in the origin of negative mobilities. The distribution function (dashed line) is strongly asymmetric around the energy minimum (band structure: plain line). Equation (4) tells us that the mobility is approximately the integral over the projection of the $\mathbf{k}$ gradient of the distribution function on the velocity divided by the total scattering rate. Applying this approximation to Fig. 5, neglecting the small contribution of the distribution function at negative $\mathbf{k}$, one sees that, from $\mathbf{k} = 0$ to the maximum of the distribution function, the $\mathbf{k}$ gradient is positive as well as the

velocity, resulting in a negative contribution to the mobility. From the maximum of the distribution function to the higher $\mathbf{k}$, the $\mathbf{k}$ gradient changes its sign, yielding a positive contribution to the mobility. Since phonon scattering is included in our scattering model, the total scattering rate $(W_{\text{tot}}(\mathbf{k}))$ grows faster than the velocity, which is approximately proportional to the square root of the energy. Therefore, as the $\mathbf{k}$ gradient of the distribution function has approximately the same magnitude on both sides of the maximum, the left part will dominate, producing a local negative mobility. The mechanism leading to the negative mobilities can be summarized, as illustrated in Fig. 6. In the following explanation, the NIN structure is taken as reference (see Fig. 1). The electrons in the highly doped emitter, which are almost in equilibrium (plain line), drift toward the NI junction. There, a small potential barrier breaks the symmetry of the distribution function and the mobility rises because the doping concentration abruptly decreases. From there on, the electrons enter a linearly increasing electric field. This field is too low for allowing quasi-ballistic motion on the next 7 nm (dotted line) but, then, becomes high enough such that the electrons have no time to scatter on the next 6 nm. Electrons with positive velocities simply fly in the transport direction in real space as well as in momentum space. Thus, the distribution function is strongly shifted from its equilibrium position

(dashed line), leading to negative mobilities for the reasons explained before. This leads to the most physical interpretation of negative mobilities: A local negative mobility is an indication of strong quasi-ballistic transport and means, contrarily to the intuition associated with a negative mobility, that almost all electrons are flying in the same direction, reverse to the electric field (i.e., the "natural" direction). Furthermore, a negative mobility does not violate any physical principle like, e.g., current continuity. Finally, it seems that negative mobilities appear only in regions where the diffusion term dominates over the drift term.

V. CONCLUSION

Although the dimensions of MOSFETs are incessantly shrinking, the grids used in TCAD simulations are almost scale-invariant leading to an increase of the spatial resolution. This, in turn, leads to the need for accurate TPs on such small scales. Computing TPs, using the exact method presented in [8], can lead to negative mobilities in certain parts of the device which must be correctly interpreted and which by no means invalidate the general definition of the mobility given in [1]. In this brief, we have demonstrated that a local negative mobility must be physically interpreted as an indication of strong quasi-ballistic transport.

APPENDIX

In the case of a parabolic-band structure, the ij component of the mobility reduces to

$$\mu_{ij} = -\int_K \frac{v_i(\mathbf{k})\partial_{\mathbf{k_j}}\mathbf{f}(\mathbf{k})}{\mathbf{W}_{\text{tot}}(\mathbf{r},\mathbf{k})} d^3k \qquad (16)$$

where $W_{\text{tot}}(\mathbf{r},\mathbf{k})$ is the total scattering rate out of the state $\mathbf{k}$ at position $\mathbf{r}$ [1]. Comparing (4) with (16), one finds that, in that case, the relaxation time $\tau(\mathbf{r},\mathbf{k})$ is nothing but

$$\tau(\mathbf{r},\mathbf{k}) := 1/W_{\text{tot}}(\mathbf{r},\mathbf{k}) \qquad (17)$$

i.e., the inverse of the total scattering rate. In well-known cases, the general definition (8) reduces to the more familiar expression (4). Therefore, definition (8) must be considered as a generalization of the popular definition (4) and not as something completely new and/or exotic.

REFERENCES

[1] S. C. Brugger, "Computation of semiconductor properties using moments of the inverse scattering operator of the Boltzmann equation," Ph.D. dissertation, Hartung Gorre Verlag, Konstanz, Germany, 2005. ETH Zürich [Online]. Available: http://e-collection.ethbib.ethz.ch/show?type=diss&nr=16376

[2] S. C. Brugger and A. Schenk, "New one-particle Monte Carlo method for nanoscale device simulation," in *Proc. Tech. NSTI Nanotechnology Conf. and Trade Show*, Boston, MA, 2006, vol. 1, pp. 673–676.

[3] J. Y. Choi and J. G. Fossum, "Analysis and control of floating-body bipolar effects in fully depleted submicrometer SOI MOSFETs," *IEEE Trans. Electron Devices*, vol. 38, no. 6, pp. 1384–1391, Jun. 1991.

[4] H. Gan and T. Tang, "A new method for extracting carrier mobility from Monte Carlo device simulation," *IEEE Trans. Electron Devices*, vol. 48, no. 2, pp. 399–401, Feb. 2001.

[5] S. Bandyopadhyay, M. Klausmeierbrown, C. Maziar, S. Datta, and M. Lundstrom, "A rigorous technique to couple Monte-Carlo and drift-diffusion models for computationally efficient device simulation," *IEEE Trans. Electron Devices*, vol. ED-34, no. 2, pp. 392–399, Feb. 1987.

[6] C. Jungemann, private communication, 2006.

[7] O. Madelung, *Introduction to Solid-State Theory*, ser. Solid-State Sciences. New York: Springer-Verlag, 1996.

[8] S. C. Brugger, A. Schenk, and W. Fichtner, "Moments of the inverse scattering operator of the Boltzmann equation: Theory and applications," *SIAM J. Appl. Math.*, vol. 66, no. 4, pp. 1209–1226, 2006. [Online]. Available: http://link.aip.org/link/?SMM/66/1209/1

[9] D. Neamen, *Semiconductor Physics and Devices*. New York: McGraw-Hill, 2003.

[10] C. Jungemann and B. Meinerzhagen, *Hierarchical Device Simulation. The Monte-Carlo Perspective*, ser. Computational Microelectronics. New York: Springer-Verlag, 2003.

[11] M. Cohen and J. Chelikowsky, "Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors," *Phys. Rev. B, Condens. Matter*, vol. 14, no. 2, pp. 556–582, Jul. 1976.

[12] M. Cohen and J. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, 2nd ed. New York: Springer-Verlag, 1989.

[13] C. Jacoboni and L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Rev. Modern Phys.*, vol. 55, no. 3, pp. 645–705, 1983.

[14] F. M. Bufler, A. Schenk, and W. Fichtner, "Efficient Monte Carlo device modeling," *IEEE Trans. Electron Devices*, vol. 47, no. 10, pp. 1891–1897, Oct. 2000.

[15] E. Cartier, M. Fischetti, E. Eklund, and F. McFeely, "Impact ionization in silicon," *Appl. Phys. Lett.*, vol. 62, no. 25, pp. 3339–3341, Jun. 1993.

Simon C. Brugger received the Diploma in physics and the Ph.D. degree in technical sciences from the Swiss Federal Institute of Technology, Zürich, Switzerland, in 2000 and 2005, respectively.

Since then, he has been with the Integrated System Laboratory in the Technology Computer Aided Design (TCAD) research group. His professional interests include semiconductor physics, applied mathematics, and Monte Carlo simulation.

Andreas Schenk received the Dipl.Phys. and Ph.D. degrees from Humboldt University, Berlin, Germany, in 1981 and 1987, respectively.

From 1987 to 1991, he was working on various aspects of the physics and simulation of optoelectronic devices. In 1991, he was with the Integrated Systems Laboratory (ISL), ETH Zürich, as a Senior Research/Teaching Assistant, where he qualified to give lectures at university in 1997 for "Physics and Modeling of Microelectronic Devices." In 2004, he became a Titular Professor at the ISL. His main activities are in the physics-based modeling of micro- and nanoelectronic devices: Quantum effects in ultra-short transistors, physics and simulation of single-electron devices, many-body effects, contacts, heterojunctions, degradation effects, as well as transport processes like resonant and sequential tunneling, generation–recombination, and noise. He is currently heading the Device Physics Group at ISL. He authored and coauthored two books and 125 papers.