A Wigner Function-Based Quantum Ensemble Monte Carlo Study of a Resonant Tunneling Diode

L. Shifren, C. Ringhofer, and D. K. Ferry

Abstract—We present results of resonant tunneling diode operation achieved from a particle-based quantum ensemble Monte Carlo (EMC) simulation that is based on the Wigner distribution function (WDF). Methods of including the Wigner potential into the EMC, to incorporate natural quantum phenomena, via a particle property we call the affinity are discussed. Dissipation is included via normal Monte Carlo procedures and the solution is coupled to a Poisson solver to achieve fully selfconsistent results.

I. INTRODUCTION

AS CURRENT device technologies quickly approach the scales whereby quantum effects due to the strong confinement of carriers and direct source-drain tunneling will begin to dominate [1], [2], new simulation techniques are required in order to fully understand the physics behind the technology operation. Of all the simulation methods currently employed, ensemble Monte Carlo (EMC) has always been the most vigorous and trusted method for device simulation, as it is proven to be reliable and predictive [3]. However, as EMC relies on the particle nature of the electron, quantum effects associated with the wave-like nature of the electron cannot be fully incorporated into the simulation. In order to resolve quantum mechanical effects, the wave-like nature of the electron needs to be incorporated into the EMC. To accomplish this effect, attention is turned to the quantum mechanical Wigner distribution function (WDF), which has found success in modeling resonant tunneling diodes (RTD) [4]–[7].

The EMC is a stochastic method used to solve the Boltzmann transport equation (BTE). The similarities between the BTE and the Wigner transport equation (WTE) [8], [9], along with the fact that both utilize localized distributions, naturally leads to the conclusion that EMC should be a valid method for solving the WTE. The main difference between the BTE and the WTE is the nonlocal potential term of the WTE. The BTE, being classical in nature, treats the potential as a localized force term. This localized force term is purely classical. The WTE treats the potential term nonlocally and fully incorporates quantum effects such as tunneling and correlation, and causes the WDF to have

negative parts. Therefore, unlike the distribution in the BTE, the WDF is not a normal probability function in that it can have negative values (negative probabilities have been discussed by Feynman [10]).

In order to account for the negative parts of the WDF, which cannot be accommodated in a normal EMC, we assign the particles in the EMC a new property that we term the particle affinity [11]. The affinity is a weighting given to the particle that represents its contribution to the total charge distribution of the system. The magnitude of the particle affinity is limited to be less then 1, but it can take on negative values accounting for negative probabilities. That is, the particle in the simulation may not have a value greater then a single electron, but this value may be less then 0. We return to this point later.

An additional difference between direct solutions of the WTE and EMC solutions of the BTE is that, although the physics that govern scattering in both transport equations is identical, incorporating scattering into the WTE is not an easy process and has not yet been fully achieved [12]. In contrast to this, Monte Carlo methods have been very successful in incorporating scattering into BTE simulation for years, and the physics and numerical methods behind these dissipative Monte Carlo techniques is well understood and tested. Therefore, by using Monte Carlo techniques to solve the WTE, augmented with the particle affinity, it is possible to incorporate quantum effects, such as tunneling, into a method previously unable to demonstrate such physical processes and at the same time include the scattering term of the WTE fully into the Wigner problem. Our method depends on calculating the Wigner potential exactly, and updating the electron distribution within the standard EMC to account for the nonlocal correction to the density in the system. This is achieved by directly updating the particle affinity. A simplified version of this method has already been used by us to study Gaussian tunneling through a single barrier nonself-consistently [11]–[14]. Other non-WDF methods to include quantum effects have been developed, such as the effective potential [15], [16], which account for some phenomena associated with the wave-like nature of the electron but cannot account for tunneling, correlation or interference effects [17].

The WTE is given by

L. Shifren and D. K. Ferry are with the Center for Solid State Electronics Research, Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287 USA (e-mail: ferry@asu.edu).

C. Ringhofer is with the Department of Mathematics, Arizona State University, Tempe, AZ 85287 USA.

$$\frac{df_w(x,k,t)}{dt} - \frac{df_w(x,k,t)}{dt}\bigg|_{collision}$$
$$+ \mathbf{v}\frac{df_w(x,k,t)}{dx} - \theta\left[f_w(x,k,t)\right] = 0 \quad (1)$$

where the second term accounts for scattering, the third term is the "diffusive" term and the fourth term incorporates the potential in the system, and is given by

$$\theta\left[f_w\left(x, k, t\right)\right] = \frac{1}{2\pi\hbar} \int dk' W\left(x, k'\right) f_w\left(x, k + k', t\right) \tag{2a}$$

where the kernel is

$$W\left(x, k'\right) = \int dx' \sin\left(k'x'\right)\left[V\left(x + \frac{x'}{2}\right) - V\left(x - \frac{x'}{2}\right)\right]. \tag{2b}$$

As V in (2b) is the total potential, which includes both the barrier and the selfconsistent applied potentials, calculating such a kernel is excessively difficult. Moreover, it has to be continuously updated since the selfconsistent potential is dynamic. It can be shown that, if the potential profile is of second-order or lower, (2) reduces to [12]

$$\theta\left[f_w\left(x, k, t\right)\right] = \frac{1}{\hbar} \frac{dV\left(x\right)}{dx} \frac{df\left(x, k, t\right)}{dk} \tag{3}$$

which should be recognized as the classical force term of the BTE. As the selfconsistent potential is slowly varying, it is possible to decouple the latter from the barrier potential, in the case of the RTD structure. In our approach, barrier potentials, which are static, are treated using (2) while the selfconsistent potential is included classically using (3). We have checked this in the numerical results, and find that the selfconsistent potential is quadratic or less to better than 99% in this problem. The final equation that we, therefore, solve is given as

$$\frac{df_w\left(x, k, t\right)}{dt} - \frac{df_w\left(x, k, t\right)}{dt}\bigg|_{collision} + \mathbf{v}\frac{df_w\left(x, k, t\right)}{dx}$$
$$- \frac{1}{\hbar}\frac{dV\left(x\right)}{dx}\frac{df\left(x, k, t\right)}{dk} - \theta\left[f_w\left(x, k, t\right)\right] = 0. \tag{4}$$

On close examination, it becomes evident that the first four terms of (4) are the BTE, whereby the last term is the quantum Wigner potential. This is the basis of the method we use to solve this equation.

The Monte Carlo technique is set up such that two systems are solved simultaneously. The first system is the particle system, which resembles a standard classical EMC. The second system is the wave properties of the particles—the affinity. That is, all particles in the system are treated classically as whole particles. They are scattered using normal EMC scattering techniques, and are drifted and accelerated using the standard field term caused by the Hartree potential (solution of the Poisson equation). Once the above operations have completed, the Wigner distribution function is calculated from the particle's position and affinity according to

$$f_w\left(x, k\right) \Leftrightarrow \sum_i \delta\left(x - x_i\right)\delta\left(k - k_i\right)A\left(i\right) \tag{5}$$

where k_i, x_i and $A(i)$ are the momentum, position and affinity, respectively, of the ith particle. There are two points here. By using this in the first term of (4), we see that one needs to temporally update both the classical properties (position, momentum, etc.) and the quantum properties (affinity). The first of these

updates, discussed above, is done by the normal Monte Carlo technique. The second is the affinity update. By constructing the distribution function, we are then able to utilize (2) to calculate the nonlocal Wigner potential term (NLP), *which determines the change of affinity that each particle experiences due to the quantum structure in the system*. That is, it is the last term of eqn. (4) that updates the wave-like properties of the particles through the update of their affinity. This can be summarized as follows. All particles in the system are drifted, accelerated and scattered, regardless of their affinities. The particle affinities are changed by the NLP. It is clear that all the quantum mechanics is incorporated into the method via the NLP and the variation of the particle affinity. An alternate realization of the method is to recognize that the particles themselves do not see the quantum barriers, only their affinities "see" the barrier.

As eluded to previously, this is a quantum ensemble Monte Carlo, that is, we retain the full particle nature of the EMC technique. We are able to utilize full ensemble statistics by noting that any ensemble average takes the form

$$\langle Q \rangle = \frac{\sum_i A\left(i\right)Q\left(i\right)}{\sum_i A\left(i\right)} \tag{6}$$

where Q is the quantity of interest, such as the velocity or the energy. It should be noted then if we set the particle affinity to 1 (i.e., a classical EMC), we regain the well known definition of the ensemble. The total number of electrons in the system is, therefore, given as

$$N_{electrons} = \sum_i A\left(i\right). \tag{7}$$

Because some $A(i) < 0$, we require N, the number of simulated particles, to be much larger then $N_{electrons}$. To achieve this, we define a maximal envelope (ME) which is a larger particle distribution that defines N. Here, the ME at any phase point is larger then the magnitude of f_w, as if all $A(i) = 1$. For example, the ME is any distribution such that $ME(x, k) \geq |f_w(x, k)|$. The particle density initially is then spread physically over the ME, according to its variation in positon and momentum. The initial affinity is assigned by the following procedure: if a particle has position x and momentum k, then its affinity is given value 0 if that position and momentum is not occupied under $f_w(x, k)$. That is, the number of particles in a small region $\Delta x \Delta k$ are assigned affinity 1 according to the number $f_w(x, k)\Delta x \Delta k$. If more particles are within this region, the excess particles are given affinity 0. In the equilibrium situation, the Wigner function is positive definite, so we do not have negative affinities initially. The affinities are then updated by the procedure described above. A further importance of using the ME is to ensure that enough particles are present to not only gain and lose affinity due to the NLP, but also to sample the entire phase-space domain of the WDF. As the NLP term acts nonlocally, correlation, reflection and transmission of density can occur where the WDF is zero or negative. Since the particles act as charge carriers, particles need to correctly sample the phase-space domain to correctly incorporate the nonlocal updates. In Fig. 1, we illustrate the effectiveness of this approach by plotting the tunneling

of the Gaussian wave packet through a single barrier of width 3 nm and height 0.3 eV. Clearly observable are the reflected and transmitted parts, as well as the correlation that remains in the barrier region. The total wave function, as shown in the figure, is termed a macroscopic quantum superposition [18].

One last critical aspect is the phase-space discretization. In standard EMC, discretization occurs only for the Poisson solver in real space. However, we are required to calculate the WDF, which is discretized in both position and momentum. Because of the periodicity associated with discretizing (2b), the spacing in position and momentum are related by

$$\Delta k = \frac{\pi}{N_x \Delta x} = \frac{\pi}{L} \tag{8}$$

where L is the device length and N_x is the number of grid points in space. This periodicity also leads to a maximum momentum the system should sample, which is given as

$$k_{\max} = \frac{\pi}{\Delta x}. \tag{9}$$

The choice of Δx therefore determines the largest momentum in the simulation. Δx is chosen so that it is small enough to sense quantum effects (in our case, we utilize 1 nm). Although a standard EMC sets no limits on momentum, with a choice of 1 nm for Δx, we find that the energy associated with the maximum momentum is on the order of 1.4 eV, which is sufficiently larger than any energy of interest in our system.

Boundary conditions are required such that the EMC, ME, and WDF boundary conditions are all satisfied. This is achieved by randomly distributing particles in the contact region during injection (EMC condition), then randomly distributing these particles in momentum according to the ME and finally, assigning the particle affinities based on a thermal distribution function in the contact such that charge neutrality at the contacts is met. This last step incorporates not only the WDF boundary condition, but also the need for charge neutrality. Although previous work has suggested the need for a drifted Maxwellian boundary conditions [19], we find satisfactory results using thermal distributions at the contacts. Also, absorbing boundary conditions are required when solving the WDF [20] to prevent spurious reflections. However, the EMC and ME condition on the boundary naturally includes this. Due to the absorbing nature of the boundary conditions and the need to update the ME to assure proper sampling of the entire phase-space, current is calculated from the probability current, which is known to be

$$J(x, t) = q |\Psi(x, t)|^2 v(x, t). \tag{10}$$

By using our definition, and properties of the WDF [8], it is easily shown that the current in the device becomes

$$I(t) = \frac{q\hbar}{mL} \sum_i A(i) \cdot k(i). \tag{11}$$

The device simulated here is a 1-D (in space) RTD which consists of 3 nm, 0.3 eV barriers surrounding a 5-nm quantum well. The barrier structure is centered in a 30-nm lightly doped (10^{16} cm^{-3}) spacer region that is connected to 60-nm highly doped (10^{18} cm^{-3}) drift regions on either side. The device
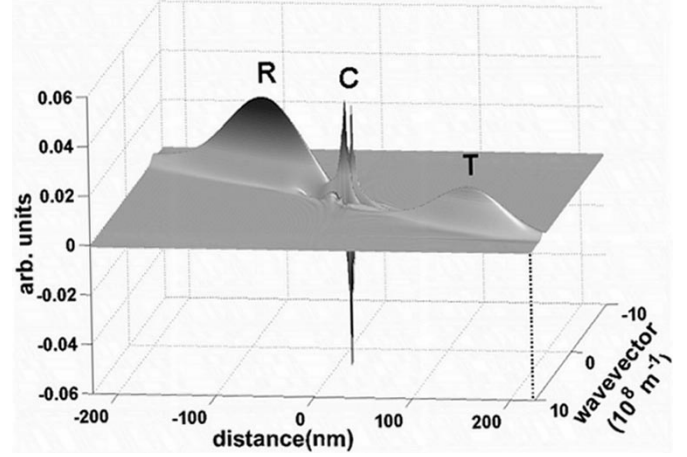

Fig. 1. The Wigner distribution function for the correlation (C), transmitted (T), and reflected (R) wave packets.

is considered to be entirely GaAs ($m^* = 0.067$), which is degenerate at these doping levels. Full Fermi–Dirac statistics (incorporated via the projection onto the spatial grid, with some coarse-graining) in the scattering are used in the calculation, which is performed at 300 K. Polar optical scattering, with both absorption and emission terms, is included using standard EMC techniques [21]. Boundary conditions are critical, as in any device simulation. Here, we ensure that current and electron number is conserved in the simulation. That is, current continuity across both contacts, and net space-charge neutrality are enforced. Selfconsistency is included by solving the Poisson equation, by a direct matrix inversion. The device is run from 0–0.5 V, in 0.025 V increments, and each bias point is found to selfconsistently evolve to steady state within 2 ps. The bias is then reduced stepwise back to 0 V to check for hysteresis. The resulting I–V characteristics are shown in Fig. 2. What is evident from this I–V characteristic is that the method correctly predicts the negative differential resistance (NDR) expected in this device. Furthermore, the peak and valley locations are approximately where expected. It is of interest that our simulations do not show any hysteresis [22]. The reason for the NDR is that as the device is biased, the conduction band (CB) in the contacts begins to sweep past the resonant level inside the well. As the CB is lined up with the resonant level in the well, tunneling is maximized as is evident from the large density inside the well as seen in Fig. 3(a). As more bias is applied, the CB rises above the resonant level inside the well and tunneling into the well is reduced as is evident from the decrease in density in the well as seen in Fig. 3(b). This cutoff in tunneling results in a large accumulation of carriers before the barrier, which is also evident in the latter figure. As more bias is applied, the barrier on the drain side of the device begins to fall below the CB of the contact in the source region and single barrier tunneling begins to dominate the device operation, which results in the increasing current in Fig. 2 above 0.3 V. Furthermore, more interesting quantum effects are seen in the device operation. In Fig. 4, we see large negative correlation due to an increase in tunneling though the quantum structure [11], [14] and also clearly see the electrons that tunnel through the barriers.
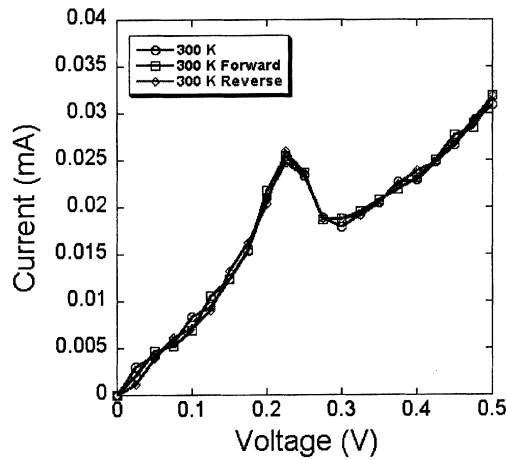
Fig. 2. Current voltage characteristics achieved by initially increases the bias from 0–0.5 V and then reducing the bias stepwise back to 0 V. An additional forward bias run is also included to demonstrate that small differences in the currents are due to EMC noise. Notice, there is no hysteresis.
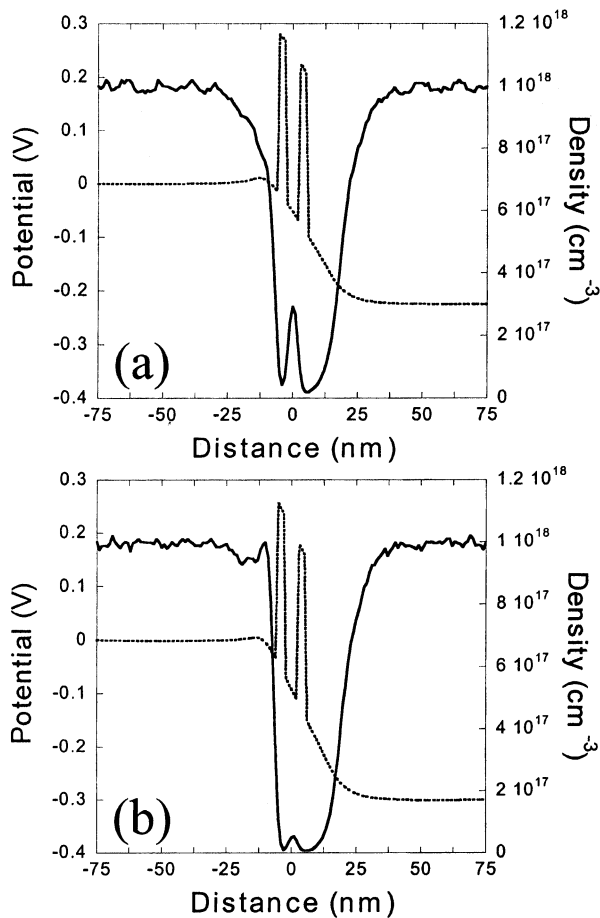

Fig. 4. The Wigner distribution function achieved selfconsistently at a bias of 0.225 (peak of I–V in Fig. 1) (A) Notice the large negative correlation which exist before the barrier due to the tunneling. (B) The density tunneling through the structure.

EMC techniques. These results allow adapting the popular EMC technique to quantum structures, thus extending the range of this important method. As simulation demands increase, and device sizes become smaller, techniques such as this QEMC will become more relevant and important. We should note a similar approach due to Kuhn and Rossi, in which a Monte Carlo simulation of the density matrix is developed [23]. The Wigner function is, of course, just a Fourier transform of the difference coordinate of the density matrix, as discussed in a subsequent publication [24], so the approaches have a common theme. However, the density matrix is real quantity, and the nonlocal potential is harder to realize in this latter approach, which was applied to optical excitations of a homogeneous system. It is not clear how this approach will work in a device format.

Fig. 3. (a) (Dashed) Selfconsistent potential and barrier potential at a bias of 0.225 V, which corresponds to the peak of the I–V curve in Fig. 1. (Solid) The associated density profile showing the a large increase of density in the well due to resonant tunneling. (b) (Dashed) Selfconsistent potential and barrier potential at a bias of 0.325 V, which corresponds to the valley of the I–V curve in Fig. 1. (Solid) The associated density profile showing the a large decrease of density in the well, as well as a large accumulation of carriers before the barrier.

This new QEMC approach clearly shows an ability to correctly simulate and incorporate quantum effects into EMC simulations. Furthermore, scattering is easily included via normal
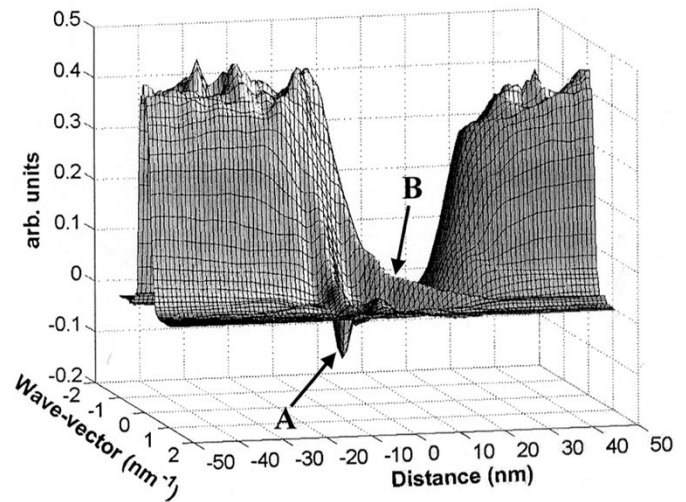
REFERENCES

[1] H. Kawaura, T. Sakamoto, T. Baba, Y. Ochiai, J. Fujita, S. Matsui, and J. Sone, "Proposal of pseudo source and drain MOSFET's for evaluating 10-nm gate MOSFETs," *Jpn. J. Appl. Phys.*, vol. 36, pp. 1569–1573, 1997.

[2] D. K. Ferry, *Granular Nanostructures*, D. K. Ferry, J. R. Barker, and C. Jacoboni, Eds. New York: Plenum, 1985, pp. 1–18.

[3] M. V. Fishcetii and S. E. Laux, "Monte-Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Phys. Rev B.*, vol. 34, no. 14, pp. 9721–9745, Nov. 1998.

[4] U. Ravaioli, M. A. Osman, W. Pötz, N. C. Kluksdahl, and D. K. Ferry, "Investigation of ballistic transport through resonant-tunneling quantum wells using the Wigner function approach," *Physica B + C*, vol. 134B, pp. 36–40, 1985.

[5] W. R. Frensley, "Wigner-function model of a resonant-tunneling semiconductor-device," *Phys. Rev. B*, vol. 36, pp. 1570–1580, 1987.

[6] N. C. Kluksdahl, W. Pötz, U. Ravaioli, and D. K. Ferry, "Wigner function study of a double quantum barrier resonant tunneling diode," *Supperlatt. Microstruct.*, vol. 3, pp. 41–45, 1987.

[7] N. C. Kluksdahl, A. M. Kriman, D. K. Ferry, and C. Ringhofer, "Self-consistent study of the resonant-tunneling diode," *Phys. Rev. B*, vol. 39, pp. 7720–7735, 1989.

[8] D. K. Ferry and H. L. Grubin, "Modeling of quantum transport in semiconductor devices," *Solid State Phys.*, vol. 49, p. 283, 1995.

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

[10] R. Feynman, "Negative probability," in *Quantum Implications, Essays in Honor of David Bohm*, B. J. Hiley and F. D. Peat, Eds. London, U.K.: Routledge & Kegan Paul, 1987, pp. 235–248.

[11] L. Shifren and D. K. Ferry, "Particle Monte Carlo simulation of Wigner function tunneling," *Phys. Lett.*, vol. A285, pp. 217–221, 2001.

[12] C. Jacoboni, A. Bertoni, P. Bordone, and R. Brunetti, "Wigner-function formulation for quantum transport in semiconductors: Theory and Monte Carlo approach," *Math. Comp. Simul.*, vol. 55, pp. 67–78, 2001.

[13] J. Garcia-Garcia, F. Martin, X. Oriols, and J. Sune, "Quantum Monte Carlo simulation of resonant tunneling diodes based on the Wigner distribution function formalism," *Appl. Phys. Lett.*, vol. 73, pp. 3539–3541, 1998.

[14] L. Shifren and D. K. Ferry, "Wigner function quantum Monte Carlo," *Phys. B*, to be published.

[15] L. Shifren, R. Akis, and D. K. Ferry, "Correspondence between quantum and classical motion: Comparing Bohmian mechanics with a smoothed effective potential approach," *Phys. Lett.*, vol. A274, pp. 75–83, 2000.

[16] R. Akis, L. Shifren, D. K. Ferry, and D. Vasileska, "The effective potential and its use in simulation," *Phys. Stat. Sol. (b)*, vol. 226, pp. 1–8, 2001.

[17] D. K. Ferry, S. M. Ramey, L. Shifren, and R. Akis, "The effective potential: The good, the bad and the ugly," J. Comp. Elec., to be published.

[18] R. B. Griffiths, *Consistent Quantum Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2002, pp. 134–136.

[19] W. R. Frensley, "Boundary conditions for open quantum systems driven far from equilibrium," *Rev. Mod. Phys.*, vol. 62, pp. 745–791, 1990.

[20] C. Ringhofer, D. K. Ferry, and N. C. Kluksdahl, "Absorbing boundary conditions for the simulation of tunneling effects in semiconductors," *Trans. Theory Stat. Phys.*, vol. 18, pp. 331–346, 1989.

[21] D. K. Ferry, *Semiconductors*. New York: Macmillan, 1991, pp. 221–223.

[22] H. L. Grubin, R. C. Buggeln, and J. P. Kreskovsky, "Boundaries, transients and multiple barrier devices," *Supperlatt. Microstruct.*, vol. 27, no. 5/6, pp. 533–537, 2000.

[23] T. Kuhn and F. Rossi, "Analysis of coherent and incoherent phenomena in photoexcited semiconductors: A Monte Carlo approach," *Phys. Rev. Lett.*, vol. 69, pp. 977–980, 1992.

[24] F. Rossi, A. di Carlo, and P. Lugli, "Microscopic theory of quantum-transport phenomena in mesoscopic systems: A Monte Carlo approach," *Phys. Rev. Lett.*, vol. 80, pp. 3348–3351, 1998.

L. Shifren, photograph and biography not available at the time of publication.

C. Ringhofer, photograph and biography not available at the time of publication

D. K. Ferry, photograph and biography not available at the time of publication