Unified particle approach to Wigner-Boltzmann transport in small semiconductor devices

M. Nedjalkov,[1,2] H. Kosina,[2] S. Selberherr,[2] C. Ringhofer,[3] and D. K. Ferry[1]
[1]*Department of Electrical Engineering, Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-1804, USA*
[2]*Institute for Microlectronics, TU-Vienna, Gusshausstrasse 27-29, E360 A-1040, Vienna, Austria*
[3]*Department of Mathematics, Arizona State University, Tempe, Arizona 85287-1804, USA*
(Received 20 March 2004; published 23 September 2004)

Small semiconductor devices can be separated into regions where the electron transport has classical character, neighboring with regions where the transport requires a quantum description. The classical transport picture is associated with Boltzmann-like particles that evolve in the phase-space defined by the wave vector and real space coordinates. The evolution consists of consecutive processes of drift over Newton trajectories and scattering by phonons. In the quantum regions, a convenient description of the transport is given by the Wigner-function formalism. The latter retains most of the basic classical notions, particularly, the concepts for phase-space and distribution function, which provide the physical averages. In this work we show that the analogy between classical and Wigner transport pictures can be even closer. A particle model is associated with the Wigner-quantum transport. Particles are associated with a sign and thus become positive and negative. The sign is the only property of the particles related to the quantum information. All other aspects of their behavior resemble Boltzmann-like particles. The sign is taken into account in the evaluation of the physical averages. The sign has a physical meaning because positive and negative particles that meet in the phase space annihilate one another. The Wigner and Boltzmann transport pictures are explained in a unified way by the processes drift, scattering, generation, and recombination of positive and negative particles. The model ensures a seamless transition between the classical and quantum regions. A stochastic method is derived and applied to simulation of resonant-tunneling diodes. Our analysis shows that the method is useful if the physical quantities do not vary over several orders of magnitude inside a device.

DOI: 10.1103/PhysRevB.70.115319 PACS number(s): 73.23.−b, 02.60.Cb, 02.70.Ss, 72.10.Bg

I. INTRODUCTION

A. The Wigner-Boltzmann-transport

Carrier transport in mesoscopic devices has been widely investigated in recent years using the Wigner function formalism.[1] A single-particle picture[2] is used, where the coherent interaction of the electron with the device structure is determined by the Wigner potential. This approach allows one to handle, in a natural way, self-consistent open-boundary systems under stationary, small signal, or transient conditions.[3] Early works investigate the theoretical and numerical properties of the coherent Wigner equation appropriate for ballistic transport.[4–8] Dissipative interactions with phonons have been approached by means of phenomenological models based on the relaxation-time approximation.[3,9–11] The Boltzmann collision operator, acting upon the Wigner distribution, has been suggested by Frensley[9] as a more general model of dissipative processes caused by phonons. Can the classical collision operator and the quantum Wigner operator reside in a common equation? The answer can be found by starting from the first-principle equation for the generalized Wigner function.[12,13] Along with the electron coordinates, the function depends on the occupation number of the phonon states in the system. Of interest is the electron, or reduced, Wigner function obtained from the generalized Wigner function by a trace over the phonon coordinates. A closed equation for the reduced Wigner function can be derived after a hierarchy of approximations, which include the weak scattering limit and assume that the phonon system is in equilibrium.[14] They concern the interaction with the phonons, while the potential operator remains exact. The phonon interaction in the resulting equation, being nonlocal in both space and time, is yet quantum. The Wigner-Boltzmann equation is obtained after a classical limit in the phonon term, leading to the Boltzmann collision operator. The hierarchy of the corresponding transport models is shown on Fig. 1.

The effects neglected by this limit can be studied from the homogeneous form of the equation for the reduced Wigner function. In this case the latter reduces to the Levinson equation,[15] or equivalently to the Barker-Ferry equation[16] with infinite electron lifetime. It should be noted that the Barker-Ferry equation can be alternatively derived[17] from the one-band semiconductor model,[18] and inclusion of a finite lifetime requires a refined set of approximations in the first-principle equation. Numerical studies[19–22] of the Levinson equation reveal quantum effects of collision broadening, retardation, and the intracollisional field effect. These effects, related to the finite duration of the collision process, have been investigated by Ferry and Barker,[16,23–26] the Modena group,[13,27–30] and others[21,31–33] for ultrafast and/or high-field transport in semiconductors and insulators, and Rossi and Kuhn[18,34,35] and others[36,37] in photoexcited semiconductors. The solutions of the Levinson equation demonstrate the establishment of the classical, energy conserving delta function for long times. Semiclassically forbidden states are occupied at early evolution times.[18,35,38] At higher times, which are

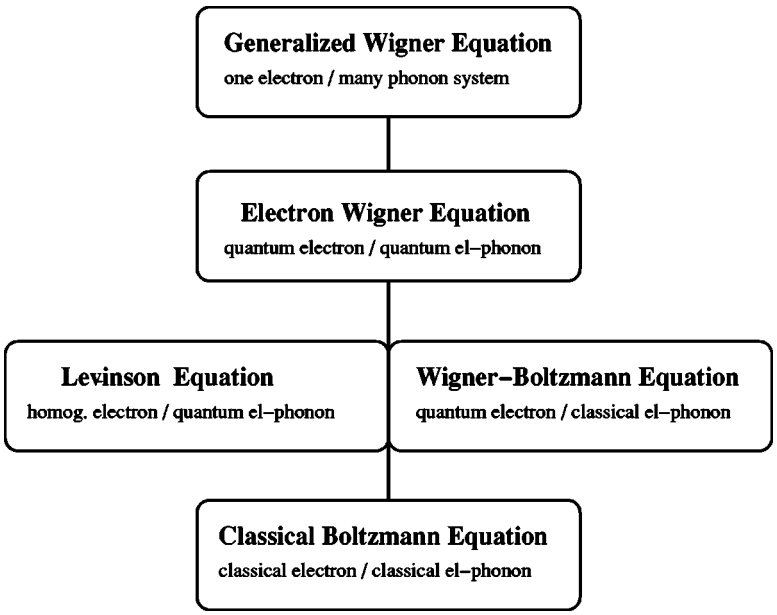

FIG. 1. Hierarchy of quantum transport models. The electron Wigner equation turns into the Levinson equation for homogeneous conditions or into the Wigner-Boltzmann equation for classical electron-phonon interaction. The Boltzmann equation is obtained with a classical limit in the Wigner potential term.

above a few hundred femtoseconds for GaAs, the Boltzmann limit dominates in the carrier evolution. A theoretical analysis supports this result. Ringhofer *et al.*[39] have derived the classical limit and the first-order correction of the equation by using a small parameter. The latter requires that the product of the time scale and the phonon frequency scale to become much larger than unity, which gives rise to coarse graining in time. Thus, for long evolution times, the quantum effects in the electron-phonon interaction can be neglected. Consequently the intracollisional field effect is not important in stationary high-field transport in semiconductors when single-valley transport is considered.[33] We note that the above considerations hold in the weak collision limit, where the next interaction begins well after the completion of the current one.

We conclude that an inclusion of the Boltzmann operator in the Wigner equation requires that the dwell time of the carriers inside the device, and hence the device itself, must be sufficiently large. On the contrary, the application of the Wigner potential operator is reasonable for small device domains, where the potential changes over a region are comparable with the coherence length of the electron. These requirements are not contradictory, as the actual device is composed of the active quantum domain attached to large contact regions.

B. Particle models

Particle models are developed for computation of physical quantities by stochastic or deterministic approaches in the framework of different quantum-kinetic theories. It has been recognized that these models, primarily introduced for numerical purposes, can be used to interpret and explain pure quantum phenomena, such as tunneling and interference. Numerical particle models are based on the convenient ideas of the classical Boltzmann model. The most simple version of the latter is built up on the free-electron quasiparticle concepts of effective mass and energy dispersion. Expansion of the physical model with respect to the band structure,[40–42] scattering with lattice imperfections,[41,43–45] Pauli exclusion principle,[46,47] and particle-particle[48] scattering retain the picture of developing particles. Below we summarize these models, starting with the direct application of the classical picture.

The smoothed effective potential approach[49] utilizes classical particles to account for quantum mechanical size quantization effects. The effective potential is a smoothing of the real classical potential due to the finite size of the electron wave packet. It has been shown that the classical trajectories resulting from the effective potential have important details in common with the corresponding Bohm trajectories.[50] A further generalization of the approach replaces the action of the Hamiltonian on the wave function by the action of a classical Hamiltonian on particles with an appropriately modified potential. A set of coupled equations is obtained for the inhomogeneous equilibrium distribution function in the device and its first-order correction. The effective potential, defined in terms of a pseudodifferential operator acting on the device potential, becomes also a function of the momenta of the classical particles.[51,52]

Ultrafast phenomena in photoexcited semiconductors are described by a set of coupled equations, where the distributions of the electrons and holes and the interband polarization are treated as independent dynamical variables. The chosen representation utilizes density matrices defined in the $(\mathbf{k},\mathbf{k}')$ wave-vector space. If interaction processes are treated on a semiclassical level, so that all transition functions become positive, the set of equations has the structure of rate equations that can be solved with a generalized ensemble Monte Carlo (EMC) scheme.[53] The fact that a particle model is associated with the evolution of the inter-band polarization, a complex quantity responsible for the coherence in the photogeneration processes, shows that: (i) the EMC method has evolved beyond the understanding of the method as a computer experiment that emulates natural processes; and (ii) numerical particles are introduced with the purpose to

model the dynamics of a quantity with no particle analogy. Furthermore, the positiveness of the transition functions is not a necessary condition for a Monte Carlo approach. It has been shown that the action of the Wigner potential, which is an antisymmetric quantity, gives rise to a Markov process, which can be regarded as a scattering of a particle between consecutive points in the phase-space.[54]

In Wigner representation, numerical particles are pointlike objects, whose temporal evolution determines trajectories in the $(\mathbf{r},\mathbf{k})$ phase space. Wigner trajectories can be defined by modified Hamilton equations, formulated with the help of a quantum force.[55] The latter is manifestly nonlocal in space and expressed through the Wigner potential, the Wigner function, and its derivative with respect to the momentum coordinate.[56] Wigner trajectories provide a pictorial explanation of the evolution of the quantum system and, in particular, nicely illustrate tunneling processes.[56,57] When it is possible to define a quantum force, Wigner trajectories maintain the values of the Wigner function during the time evolution and thus satisfy the Liouville theorem. However the quantum force has singularities at the points where the momentum derivative of the Wigner function becomes zero. At these points, trajectories can be created or destroyed.[58] In general, Wigner trajectories remain an auxiliary tool for modeling of quantum transport, unless the Wigner function in the quantum force term is assumed to be known. An appropriate approximation for a nearly equilibrium system is a displaced Maxwell-Boltzmann distribution function. It is easy to see that such an assumption corresponds to the zeroth-order correction in the effective potential approach. In this case the quantum force is defined everywhere except at the phase-space origin, and gives rise to an effective lowering of the peaks of the potential barriers.[59] The increase of the particle flow observed through the barriers is associated with tunneling processes.

The above considerations show that the price for a deviation from classical trajectories is high. Approaches utilizing zero force, as in the original formulation of the Wigner equation, or a local force,[60] which can be extracted from the Wigner operator enjoy the properties of the Liouville theorem. The conservation of the phase volume during the particle evolution is a key property for the method proposed in this work.

The integral form of the Wigner equation has been used to introduce a quantum transport model based on Wigner paths.[30,61] It has been shown that a ballistic evolution of a δ-like contribution to the Wigner function carries its value following a classical trajectory.[27] The action of the Wigner potential operator is interpreted as scattering, which, along with the scattering by the phonons, links pieces of classical trajectories to Wigner paths. We note that, in this model, the phonon interaction is treated fully quantum mechanically,[13] according to the first-principle equation. That is, the scattering with phonons begins with exchange of half of the phonon momentum and completes after a finite time. During this time, an arbitrary number of interactions with other phonons can be initiated and/or completed. In comparison, Levinson's equation considers a single interaction with finite duration while Boltzmann scattering is instantaneous, so that the trajectory changes with the full phonon momentum. During the evolution particles accumulate a numerical quantity called weight, which carries the quantum information for the system. The weight is taken into account in the computation of the physical averages. This model should be regarded as a theoretical achievement, as the practical implementation involves enormous computational burden.[13]

An operator-splitting method applied to the coherent Wigner equation gives rise to a picture where an ensemble of particles drift over classical trajectories. The quantum information is carried by a numerical quantity called affinity.[62–64] The method consists of consecutive steps of drift over the trajectories followed by an update of affinity from the Wigner potential. The approach is numerically tractable and shows an excellent agreement with deterministic solutions of the problem.[65] Phonon interaction and coupling to a Poisson solver can be included in a straightforward way,[66] so that the approach appears as a generalization of the Ensemble Monte Carlo method for solving the Wigner-Boltzmann equation. Recently a finite collision duration has been included.[67] In this approach the solution is determined by the evolution of an initial condition.

We propose an approach where the Wigner equation, with a Boltzmann scattering term, is interpreted as a Boltzmann equation with a generation term. The interaction with the Wigner potential gives rise to generation of particle pairs with opposite sign. The sign is the basic property that outlines the introduced numerical particles from classical quasiparticles. It is an important property because positive and negative particles annihilate one another. The negative values of the Wigner function in certain phase-space regions can be explained in a natural way by the accumulation of negative particles in these regions. The Wigner-Boltzmann transport process corresponds to drift, scattering, generation, and annihilation of these particles. The process unifies classical and quantum regions within a single-transport model. Although the model proposed is valid for general transport conditions, it is especially convenient for stationary transport, determined by the boundary conditions of the problem.

The cogent analogy between the concepts holds in numerical and quasiparticle approaches. Analytical models decipher complicated interactions in solids in terms of quasiparticles. Some of the real particle characteristics, such as mass and energy, can be renormalized and/or extra properties, such as lifetime, can be assigned. A canonical transform of a given analytical model can give rise to a new type of quasiparticle, so that even their total number may not be conserved.[68] Numerical models are developed to solve and/or explain given transport theory, which is built upon a corresponding quasiparticle model. The numerical particles can modify some properties of the underlying quasiparticles[69] and add novel features, such as weight or sign. Within a given numerical model, the properties of the numerical particles depend on the chosen algorithm so that their statistics and total number may not be conserved.

The paper is organized as follows. In Sec. II, we consider three equivalent formulations of the Wigner-Boltzmann equation. The adjoint integral form of the equation is used to expand the averaged value of any physical quantity of interest into a series. In Sec. III, the series is analyzed in terms of random variables. A chain of conditional probability densi-

ties associated with stochastic events of particle evolution is obtained. Section IV unifies these events into a particle transport model. Here the particle picture is explored for different transport regimes. The transport behavior of classical carriers is recovered for the case of the Boltzmann equation and the ergodicity of the process is proven. In the coherent case, governed by the Wigner equation, the transport is characterized by generation of positive and negative particles. In the general case of Wigner-Boltzmann transport, the classical and coherent models can be constructively unified. Section V discusses the numerical aspects of the model. An algorithm, suitable for Monte Carlo simulations of Wigner-Boltzmann transport, is derived. In Sec. VI, simulation results for the resonant-tunneling diode are presented and discussed.

II. THE WIGNER-BOLTZMANN EQUATION

We consider the case of stationary transport, where the physical conditions imposed on the boundaries determine the device behavior. The Wigner phase-space is composed by the real space coordinate $\mathbf{r}$ and the wave vector $\mathbf{k}$. We emphasize that $\mathbf{r}$ and $\mathbf{k}$ are independent variables, not adjoint by a Fourier transform, and thus not linked by the uncertainty relation.[9] The device exchanges carriers with two or more reservoirs through the contacts denoted by b. Open-system boundary conditions are provided by the Fermi-Dirac distribution functions $f_b(\mathbf{r}_b,\mathbf{k})$ in the contacts.[9] The solution $f_w(\mathbf{r},\mathbf{k})$ of the Wigner-Boltzmann equation is used to obtain the average values of all physical quantities of interest. The average value $\langle A\rangle$ of a generic physical quantity $A(\mathbf{r},\mathbf{k})$, such as carrier velocity and density, is given by the inner product (f_w,A)

$$\langle A\rangle = \int_D d\mathbf{r} \int d\mathbf{k} f_w(\mathbf{r},\mathbf{k})A(\mathbf{r},\mathbf{k}) = (f_w,A), \tag{1}$$

where D is the device domain. Equation (1) asserts that, in order to evaluate the averaged value of interest $\langle A\rangle$, one needs to know the solution f_w inside the device. An alternative expression for the mean value $\langle A\rangle$ can be found from the adjoint integral form of the WB equation. The derivation of this expression begins with the integrodifferential form of the equation.

A. Integrodifferential form

The stationary form of the equation states that the action of the differential Liouville operator is the joint action of the Wigner and Boltzmann integral operators on the solution f_w

$$\mathbf{v}(\mathbf{k})\cdot\nabla_{\mathbf{r}} f_w(\mathbf{r},\mathbf{k}) = \int d\mathbf{k}' V_w(\mathbf{r},\mathbf{k}'-\mathbf{k}) f_w(\mathbf{r},\mathbf{k}) + \int d\mathbf{k}' f_w(\mathbf{r},\mathbf{k}') S(\mathbf{k}',\mathbf{k}) - f_w(\mathbf{r},\mathbf{k})\lambda(\mathbf{k}). \tag{2}$$

Here $\mathbf{v}(\mathbf{k})=\hbar\mathbf{k}/m$ and m are the quasiparticle velocity and effective mass, respectively. $S(\mathbf{k}',\mathbf{k})$ is the rate for scattering from state $(\mathbf{r},\mathbf{k}')$ to state $(\mathbf{r},\mathbf{k})$ due to phonon interaction. $\lambda(\mathbf{k})=\int S(\mathbf{k},\mathbf{k}')d\mathbf{k}'$ is the phonon out-scattering rate. The Wigner potential at point $\mathbf{r}$ is defined by the Fourier transform of the central difference of the device potential V around $\mathbf{r}$

$$V_w(\mathbf{r},\mathbf{k}) = \frac{1}{i\hbar(2\pi)^3}\int d\mathbf{s}\ \exp(-i\mathbf{k}\cdot\mathbf{s}) \times\left(V\left(\mathbf{r}-\frac{\mathbf{s}}{2}\right)-V\left(\mathbf{r}+\frac{\mathbf{s}}{2}\right)\right). \tag{3}$$

The Liouville operator in (2), which for time-dependent problems must be augmented by a time derivative, corresponds to an unaccelerated free-streaming motion. The operator can be augmented by a force term, due to the following property of (3).[70] The Fourier transform of $\mathbf{F}(\mathbf{r})\cdot\mathbf{s}$, where $\mathbf{F}$ is an arbitrary function, gives rise to the term $V_w^F(\mathbf{r},\mathbf{k}'-\mathbf{k})=-F(\mathbf{r})\cdot\nabla_{\hbar\mathbf{k}}\delta(\mathbf{k}'-\mathbf{k})$. The latter, inserted in (2) and transferred to the left-hand side of the equation, completes with the Liouville operator, a force term. The definition of the Wigner potential is modified according to $V_w'=V_w-V_w^F$, which results in an additional term $-\mathbf{F}(\mathbf{r})\cdot\mathbf{s}$ in the brackets of (3). For potentials up to quadratic, $\mathbf{F}$ can be chosen to compensate the potential difference in the brackets of (3) and has a meaning of electrical force, $\mathbf{F}=-\nabla V$. In this case, $V_w'=0$ and (2) becomes the classical Boltzmann equation. For general potentials F can be selected from physical considerations[66,70] In this case, the Wigner potential V_w' must be understood as a generalized function.

We utilize definitions (2) and (3), where the characteristics of the Liouville operator are Newton trajectories without the acceleration term

$$\mathbf{r}(t)=\mathbf{r}+\mathbf{v}(\mathbf{k})t, \quad \mathbf{k}(t)=\mathbf{k}. \tag{4}$$

We note that the notions derived below do not make use of the particular form of (4) and thus remain valid for the modified definitions. As discussed in Appendix A, a stationary trajectory $(\mathbf{r}(t),\mathbf{k}(t))$ can always be initialized by the phase-space point $(\mathbf{r},\mathbf{k})$ at time 0. Equation (2) can be reformulated by a decomposition of the antisymmetric Wigner potential into two complementary parts as follows:

$$V_w(\mathbf{r},\mathbf{k}) = V_w^+(\mathbf{r},\mathbf{k}) - V_w^+(\mathbf{r},-\mathbf{k}), \quad \gamma(\mathbf{r}) = \int d\mathbf{k} V_w^+(\mathbf{r},\mathbf{k}).$$

The function V_w^+ equals V_w if the Wigner potential is positive and is zero otherwise: $V_w^+=V_w\xi(V_w)$ with ξ the Heaviside step function. The meaning of the function γ is discussed in the next section. By adding $\gamma(\mathbf{r})f_w(\mathbf{r},\mathbf{k})$ to both sides of (2) the equation becomes

$$[\mathbf{v}(\mathbf{k})\cdot\nabla_{\mathbf{r}}+\mu(\mathbf{r},\mathbf{k})]f_w(\mathbf{r},\mathbf{k}) = \int d\mathbf{k}'\Gamma(\mathbf{r},\mathbf{k},\mathbf{k}')f_w(\mathbf{r},\mathbf{k}'), \tag{5}$$

$$\Gamma(\mathbf{r},\mathbf{k}',\mathbf{k}) = V_w^+(\mathbf{r},\mathbf{k}'-\mathbf{k}) - V_w^+(\mathbf{r},\mathbf{k}-\mathbf{k}') + S(\mathbf{k}',\mathbf{k}) + \gamma(\mathbf{k})\delta(\mathbf{k}-\mathbf{k}'), \tag{6}$$

$$\mu(\mathbf{r},\mathbf{k}) = \lambda(\mathbf{k}) + \gamma(\mathbf{r}). \tag{7}$$

In Appendix B, Eq. (5) is transformed into an integral equation with the help of (4).

B. Integral form

The integral form of the WBE is formulated by using of the backward trajectory $(\mathbf{r},(t'),\mathbf{k}(t'))$, initialized by $(\mathbf{r},\mathbf{k})$

$$f(\mathbf{r},\mathbf{k}) = \int d\mathbf{k}' \int_{t_b^-}^{0} f(\mathbf{r}(t'),\mathbf{k}')\Gamma(\mathbf{r}(t'),\mathbf{k}',\mathbf{k}(t')) \times \exp\left(-\int_{t'}^{0} \mu(\mathbf{r}(y),\mathbf{k}(y))dy\right) + f_0(\mathbf{r},\mathbf{k}),$$

$$f_0(\mathbf{r},\mathbf{k}) = f_b(\mathbf{r}(t_b^-),\mathbf{k}(t_b^-))\exp\left(-\int_{t_b^-}^{0} \mu(\mathbf{r}(y),\mathbf{k}(y))dy\right);$$

$$t_b^- = t_b^-(\mathbf{k},\mathbf{r}). \tag{8}$$

The equation can be understood in analogy with the integral form of the Boltzmann equation. The latter is obtained by formally setting the Wigner potential to zero. Then, the exponent in (8) becomes the probability for a particle to drift without scattering by phonons during the time interval $(t',0)$ on the proper trajectory $\beta(\mathbf{r},\mathbf{k})$, which arrives at $(\mathbf{r},\mathbf{k})$. There are two contributions to the value of f in the point $(\mathbf{r},\mathbf{k})$. f_0 is the value of the boundary function f_b, which survives on β despite the action of the phonons. The other term gives cumulative contributions from previous times t': values of f located at $\mathbf{r}(t')$ scatter, according to S, from everywhere in the wave-vector space to the proper $\mathbf{k}(t')$. The particular values fS are further multiplied by the exponent to filter out the part that is scattered out of β by the phonons. This picture can be maintained when the Wigner potential is switched on. Then V_w^+ in (6) has a clear meaning of scattering due to the Wigner potential. The function γ can be interpreted as an out-scattering rate due to the Wigner potential in strict analogy with the phonon out-scattering rate λ. Then $\gamma\delta$ becomes a self-scattering function. The major difference between S and Γ comes from the fact that while the former is strictly non-negative, there is a minus sign in (6). This sign precludes a direct probabilistic treatment of the equation in terms of classical particles.

After this step, the boundary conditions f_0 appear explicitely in (8). f_0, along with the solution g of an equation adjoint to (8), give rise to the desired expression for the physical averages.

C. The adjoint equation

Equation (8) can be formally written in the standard form of a Fredholm integral equation of the second kind

$$f(\mathbf{r},\mathbf{k}) = \int d\mathbf{r}' \int d\mathbf{k}' f(\mathbf{r}',\mathbf{k}')K(\mathbf{r}',\mathbf{k}',\mathbf{r},\mathbf{k}) + f_0(\mathbf{r},\mathbf{k}). \tag{9}$$

The kernel K has been augmented to account for the $\mathbf{r}'$ integration by a spatial delta function

$$K(\mathbf{r}',\mathbf{k}',\mathbf{r},\mathbf{k}) = \int_{-\infty}^{0} dt'\Gamma(\mathbf{r}',\mathbf{k}',\mathbf{k}(t')) \times \exp\left(-\int_{t'}^{0} \mu(\mathbf{r}(y),\mathbf{k}(y))dy\right)\delta[\mathbf{r}' - \mathbf{r}(t')]\theta_D(\mathbf{r}'). \tag{10}$$

The indicator function of the simulation domain θ_D ensures the proper lower bound $t_b^-(\mathbf{r},\mathbf{k})$ of the time integral. The adjoint equation has the same kernel as (9) but the integration is carried out over the unprimed variables. The free term is chosen to be the physical quantity of interest A

$$g(\mathbf{r}',\mathbf{k}') = \int d\mathbf{r} \int d\mathbf{k} K(\mathbf{r}',\mathbf{k}',\mathbf{r},\mathbf{k})g(\mathbf{r},\mathbf{k}) + A(\mathbf{r}',\mathbf{k}'). \tag{11}$$

The solution g depends on the free term A, which is not explicitly written for simplicity of the notations. Equation (11) assumes a backward parametrization of the trajectories. Forward trajectories are introduced by first changing the integration variables from $\mathbf{r}$, $\mathbf{k}$ to $\mathbf{r}''=\mathbf{r}(t')$, $\mathbf{k}''=\mathbf{k}(t')$ back in time over the trajectory initialized by $\mathbf{r}$, $\mathbf{k}$. Applying (A3) and (A4), the adjoint equation (11) is obtained in forward parametrization. The integration on $\mathbf{r}''$ can be achieved using the delta function in K. A replacement of $\mathbf{k}''$ by $\mathbf{k}$ and $-t'$ by t gives rise to the following compact form:

$$g(\mathbf{r}',\mathbf{k}') = \int d\mathbf{k} \int_0^{\infty} dt\theta_D(\mathbf{r}')\Gamma(\mathbf{r}',\mathbf{k}',\mathbf{k}) \times \exp\left(-\int_0^t \mu(\mathbf{r}'(y),\mathbf{k}(y))dy\right)g(\mathbf{r}'(t),\mathbf{k}(t)) + A[\mathbf{r}',\mathbf{k}']. \tag{12}$$

Here, $(\mathbf{r}'(t),\mathbf{k}(t))$ is a forward trajectory initialized by $(\mathbf{r}',\mathbf{k})$. The equation has the desired property that integration is carried out over final states and that the time variable is positive. The solution of (12) can be expressed as a series, obtained by an iterative replacement of the equation into itself. It is convenient to consider first the series for the forward equation (11), which gives rise to (12) after the spatial integration. By introducing $Q=(\mathbf{r},\mathbf{k})$ the equation is written formally as

$$g(Q) = \int dQ'K(Q,Q')g(Q') + A(Q).$$

The solution is expanded into a Neumann series

$$g(Q) = \int dQ'\left[\delta(Q-Q') + \sum_{n=1}^{\infty} K^n(Q,Q')\right]A(Q') = (I-K)^{-1}A, \tag{13}$$

where I is the identity operator and $K^n(Q,Q') = \int dQ_1 K(Q,Q_1)K^{n-1}(Q_1,Q')$. The iterative series of (12) is equivalent to (13), but with the difference that all space integrations in the consecutive terms are performed. The space

integration leading to (12) links the independent variables, which makes it impossible to formulate a self-contained recursive relation for the expansion of this equation. It is important that the consecutive terms of the series for (12) are equal to the corresponding terms in (13),

$$g(Q) = \sum_{n=0}^{\infty} K^n A = (I-K)^{-1} A. \tag{14}$$

Denoting the reduced kernel in (12) by $K(\mathbf{r}',\mathbf{k}',\mathbf{k},t)$, we write explicitly the second term in (14),

$$K^{(2)}A = \int dt_1 \int d\mathbf{k}_1 \int dt_2 \int d\mathbf{k}_2 K(\mathbf{r}',\mathbf{k}',\mathbf{k}_1,t_1) \times K(\mathbf{r}_1(t_1),\mathbf{k}_1(t_1),\mathbf{k}_2,t_2) A(\mathbf{r}_2(t_2),\mathbf{k}_2(t_2)). \tag{15}$$

Equation (15) outlines the evolution of the space coordinate $\mathbf{r}$, which is now a passive variable, not taking part in the integration: The trajectory $(\mathbf{r}_1(t),\mathbf{k}_1(t))$ is initialized by $(\mathbf{r}',\mathbf{k}_1)$ and $(\mathbf{r}_2(t),\mathbf{k}_2(t))$ is initialized by $(\mathbf{r}_1(t_1),\mathbf{k}_2)$.

D. Physical averages

Multiplying (9) by g and (11) by f, integrating over unprimed and primed variables respectively, and subtracting the two equations leads to the equality $(f,A)=(f_0,g)$. In this way, the average value of the physical quantity A is expressed through the boundary conditions f_0 and the solution of (12),

$$\langle A\rangle = \int_D d\mathbf{r} \int d\mathbf{k} f_b(\mathbf{r}(t_b^-),\mathbf{k}(t_b^-)) \times \exp\left(-\int_{t_b^-}^{0} \mu(\mathbf{r}(y),\mathbf{k}(y))dy\right) g(\mathbf{k},\mathbf{r}), \tag{16}$$

where t_b^- and the backward trajectory $(\mathbf{r}(t),\mathbf{k}(t))$ are determined by $(\mathbf{r},\mathbf{k})$.

Since f_b is defined only at the boundary $\partial\boldsymbol{D}$, a transformation is needed that leads from a volume to a boundary integral. A phase-space point $(\mathbf{r},\mathbf{k})$ is bijectively mapped onto $[\mathbf{k}(t_b),\mathbf{r}_b=\mathbf{r}(t_b)]$ by the boundary time t_b (Ref. 71). This implies that the transformation must replace one of the space integrals by a time integral. One can formally augment (16) by a time integral $\int dt' \delta(t'-t_b)$ in the limits $(0,-\infty)$. After a rearrangement of the integrals, shown in Appendix C, (16) is transformed into

$$\langle A\rangle = \oint_{\partial D} d\sigma(\mathbf{r}_b) \int_{P+} d\mathbf{k}_b \int_0^{\infty} dt_0 |\mathbf{v}_\perp(\mathbf{k}_b)| f_b(\mathbf{r}_b,\mathbf{k}_b) \times \exp\left(-\int_0^{t_0} \mu(\mathbf{r}_b(y),\mathbf{k}_b(y))dy\right) g(\mathbf{r}_b(t_0),\mathbf{k}_b(t_0)). \tag{17}$$

By replacing g with the iteration series (14), $\langle A\rangle$ may be expanded into the series

$$\langle A\rangle = [b,(I-K)^{-1}A] = [\mathbf{v}_\perp f_b,(I-\tilde{K})^{-1}\tilde{A}] = \sum_{i=0}^{\infty} \langle A\rangle_i. \tag{18}$$

The second term is reformulated to facilitate further analysis. The multipliers in each term of the sum $[b,(I-K)^{-1}A]$ are formally regrouped. Now $\tilde{K}$ is the repeating term in the pattern, which is obtained from K after absorbing the exponent on the left and releasing the exponent to the right for the next $\tilde{K}$. In this way A is assigned with the last exponent to become $\tilde{A}$. The zero-order term $\langle A\rangle_0=(\mathbf{v}_\perp f_b,\tilde{A})$ is given by the right-hand side of (17), with $A(\mathbf{r}_b(t_0),\mathbf{k}_b(t_0))$ in the place of g. The first term is

$$\langle A\rangle_1 = \oint_{\partial D} d\sigma(\mathbf{r}_b) \int_{P+} d\mathbf{k}_b \int_0^{\infty} dt_0 \int d\mathbf{k}_1 \times \int_0^{\infty} dt_1 |\nu_\perp(\mathbf{k}_b)| f_b(\mathbf{r}b,\mathbf{k}b) \times \exp\left(-\int_0^{t_0} \mu(\mathbf{r}_b(y),\mathbf{k}_b(y))dy\right) \times \theta_D(\mathbf{r}_b(t_0))\Gamma(\mathbf{r}_b(t_0),\mathbf{k}_b(t_0),\mathbf{k}_1) \times \exp\left(-\int_0^{t_1} \mu(\mathbf{r}_1(y),\mathbf{k}_1(y))dy\right) A(\mathbf{r}_1(t_1),\mathbf{k}_1(t_1)). \tag{19}$$

Here the trajectory $(\mathbf{r}_1(t),\mathbf{k}_1(t))$ is initialized by $(\mathbf{r}_b(t_0),\mathbf{k}_1)$. The next term, $\langle A\rangle_2$ is derived with the help of (15). Eq. (19) is augmented by integrals on $\mathbf{k}_2$ and t_2, and A is replaced with $\theta\Gamma$. The product

$$\tilde{A}_2 = \exp\left(-\int_0^{t_2} \mu(\mathbf{r}_2(y),\mathbf{k}_2(y))dy\right) A(\mathbf{r}_2(t_2),\mathbf{k}_2(t_2)), \tag{20}$$

appears at the end of the expression for $\langle A\rangle_2$ as an integrand on the time t_2. Higher order terms in (18) are derived by induction.

The series expansion (18) is the key quantity in the treatment of the boundary value problem. It proves that knowledge of the boundary distribution is sufficient to determine arbitrary volume integrals defined by (1) and, therefore, to determine $\langle A\rangle$ uniquely. Only the subspace P_+ of boundary states $\mathbf{k}_b$ having an inward-directed velocity component $\mathbf{v}_\perp$ appears in (17), and thus determines the boundary condition. The complementary part is a *priori* unknown and comes out as a result of the transport process. The series expansion (18) will be analyzed in terms of probability densities.

III. ANALYSIS OF $\langle A\rangle$

The basic particle methods used to date for simulation of semiconductor devices were originally devised by considerations where the simulation was an emulation of the physical

process: the transport picture has been used to establish the corresponding stochastic method. The link between such physically based methods and the numerical methods of Monte Carlo integration has been established later.[72–77] Here, we follow the opposite approach: the numerical Monte Carlo theory is used to propose a common particle picture of the Wigner and Boltzmann transport processes. In Appendix D we show that this approach requires all integrals in (18) to be decomposed into probability densities and random variables. A great advantage is provided by the common structure of these integrals, which are built by the boundary term b, the consecutive applications of $\tilde{K}$, and end up with the quantity $\tilde{A}$. It is then sufficient to extract from each of these three quantities the proper probability densities. As these quantities appear in (19), we focus on that equation.

A. Injection from the boundaries

The boundary term allows a simple probabilistic interpretation. For the purpose of normalization we introduce the integrals

$$j_\perp(\mathbf{r}_b) = \int_{P_+} d\mathbf{k}|\mathbf{v}_\perp(\mathbf{k})|f_b(\mathbf{k},\mathbf{r}_b), \quad \Phi = \oint_{\partial D} j_\perp(\mathbf{r})d\sigma(\mathbf{r}), \tag{21}$$

which represent the normal component of the incident-particle current density and the total incident-particle current. Then the quantity

$$p_b(\mathbf{r}_b,\mathbf{k}_b) = \frac{j_\perp(\mathbf{r}_b)}{\Phi}\frac{|\mathbf{v}_\perp(\mathbf{k}_b)|f_b(\mathbf{r}_b,\mathbf{k}_b)}{j_\perp(\mathbf{r}_b)} = p_{b1}(\mathbf{r}_b)p_{b2}(\mathbf{r}_b,\mathbf{k}_b)$$

has the proper normalization of a conditional probability density. p_b generates a phase-space point on the boundary by first selecting the position $\mathbf{r}_b$ proportional to the incident-particle current density. $\mathbf{k}_b$ is then selected according to the velocity-weighted equilibrium distribution p_{b2}. In this way, the boundary term is factorized into a product of p_b with the normalization constant Φ. The selection of the boundary point follows the classical rules used in the device MC method and is thus associated with a particle that is injected into the device.

B. Probability factors in $\tilde{K}$

$\tilde{K}$ is augmented by a multiplication and a division by μ, which gives rise to the product

$$\tilde{K}(\mathbf{r}',\mathbf{k}',\mathbf{k},t) = p_t(t,\mathbf{r}',\mathbf{k}')\,\theta_D(\mathbf{r}'(t))\frac{\Gamma(\mathbf{r}'(t),\mathbf{k}'(t),\mathbf{k})}{\mu(\mathbf{r}'(t),\mathbf{k}'(t))},$$

$$p_t = \mu(\mathbf{r}'(t),\mathbf{k}'(t))\exp\left(-\int_0^t \mu(\mathbf{r}'(y),\mathbf{k}'(y))dy\right).$$

The structure of the first term p_t is well known from the classical MC method; it is the probability for a drift without scattering provided that the scattering frequency is μ. The normalization to unity is readily proven by integration over time with the limits $(0,\infty)$. p_t generates a value of t associated with a free flight time of a particle which drifts over a piece of a Newton trajectory between the initial state $(\mathbf{r}',\mathbf{k}')$ and $(\mathbf{r}'(t),\mathbf{k}'(t))$, and which, as can be seen from what follows, has a meaning of a before-scattering state. It is used as an input in the conditional probabilities composing the remaining term $\Gamma(\mathbf{r}'(t),\mathbf{k}'(t),\mathbf{k})/\mu(\mathbf{r}'(t),\mathbf{k}'(t))$ to generate the output value of $\mathbf{k}$:

$$\frac{\Gamma(\mathbf{r}',\mathbf{k}',\mathbf{k})}{\mu(\mathbf{r}',\mathbf{k}')} = p_\lambda(\mathbf{r}',\mathbf{k}')p_{ph}(\mathbf{k}',\mathbf{k}) + p_\gamma(\mathbf{r}',\mathbf{k}') \times \left(\tfrac{1}{3}p_w^+(\mathbf{r}',\mathbf{k}' - \mathbf{k}) - \tfrac{1}{3}p_w^-(\mathbf{r}',\mathbf{k}-\mathbf{k}') + \tfrac{1}{3}p_\delta(\mathbf{k}-\mathbf{k}')\right)3 \tag{22}$$

$$p_\lambda(\mathbf{r}',\mathbf{k}') = \frac{\lambda(\mathbf{k}')}{\mu(\mathbf{r}',\mathbf{k}')}, \quad p_\gamma(\mathbf{r}',\mathbf{k}') = \frac{\gamma(\mathbf{r}')}{\mu(\mathbf{r}',\mathbf{k}')},$$

$$p_{ph}(\mathbf{k}',\mathbf{k}) = \frac{S(\mathbf{k}',\mathbf{k})}{\lambda(\mathbf{k}')}, \quad p_w^\pm(\mathbf{r}',\mathbf{k}) = \frac{V_w^+(\mathbf{r}',\mathbf{k})}{\gamma(\mathbf{r}')}.$$

Here the time argument has been omitted, and $p_w^- = p_w^+$ is introduced for convenience. According to (7), p_λ and p_γ are two complementary probabilities, which can be used to select either p_{ph} or the term in the brackets in (22). The first branch occurs with the probability p_λ, which selects the type of interaction to be scattering with phonons. The application of the probability density p_{ph} is readily understood as a generation of the phonon after-scattering state $(\mathbf{r}',\mathbf{k})$. The second branch can be interpreted as a generation of an after-scattering state due to interaction with the Wigner potential. It is comprised of the three terms enclosed in the brackets. p_2 has been introduced with the purpose of selecting which one of the three densities p_w^+, p_w^-, and p_δ generates an after-scattering state $(\mathbf{r}',\mathbf{k})$. In this way, the action of the Wigner potential is realized by a scattering generated by either of these three probability densities. They will be discussed in detail in the next section. Here, we conclude that the consecutive application of the conditional probabilities comprising $\tilde{K}$ generates a transition between $(\mathbf{r}',\mathbf{k}')$ and $(\mathbf{r}'(t),\mathbf{k})$ which is associated with a particle which undergoes a free flight followed by a scattering event.

What remains for the random variable associated with $\tilde{K}$ is the term $w\theta_D = (\pm 3)^i\theta_D$. The power i depends on the type of the interaction: $i=0$ and $w=1$ if the scattering is due to phonons. If the Wigner potential is selected as a scattering source, $i=1$ and $w=(\pm 3)$, where the minus sign applies if p_w^- in (22) is selected. The quantity w is called a weight factor. The domain indicator θ_D is unity if the particle is inside the device at the end of the free flight and is zero otherwise. $\tilde{K}$ factorizes into a product of the random variable and an evolution operator composed by conditional probability densities.

C. Recording averages

The integrand (20) can be written as follows:

$$\tilde{A} = p_t(t,\mathbf{r},\mathbf{k})\frac{A(\mathbf{r}(t),\mathbf{k}(t))}{\mu(\mathbf{r}(t),\mathbf{k}(t))}.$$

The random variable ψ_A associated with the physical quantity A is the term A/μ evaluated at the end of the free flight. It must be noted that, as we are interested in physical averages in a given region Ω inside the device domain, A contains implicitly the indicator θ_Ω of that region. If the end point of the free flight is outside Ω, the random variable is zero. Another way to express the random variable ψ_A can be obtained by integration by parts of the t integral

$$\tilde{A} = p_t(t,\mathbf{r},\mathbf{k})\int_0^t dy A(\mathbf{r}(y),\mathbf{k}(y)). \quad (23)$$

The value of ψ_A is identified as the path integral over y. Actually, due to the indicator θ_Ω, only the part of the path belonging to Ω contributes to the integral. The two functionals of A are known in the classical Single-Particle MC method as synchronous ensemble and time integration techniques.[78]

So far we are ready to state the stochastic approach for evaluation of (19). Numerical trajectories are built up with the help of p_b, the conditional probabilities identified from $\tilde{K}$ and the probability p_t. The random variable $\psi_1 = \Phi\theta(\pm 3)^i\psi_A$ is calculated for each trajectory. The sample mean (D2) over N trajectories estimates $\langle A\rangle_1$. A generalization for the nth term in (18) is straightforward. Numerical trajectories are built up with the help of p_b, n consecutive iterations of the conditional probabilities identified from $\tilde{K}$, and the probability p_t. The corresponding random variable ψ_n is given by the product

$$\psi_n = \Phi\prod_{k=1}^{n}\theta_{D_k}(\pm 3)^{i_k}\psi_A = \Phi W_n\psi_A, \quad \psi_{\langle A\rangle} = \sum_n \psi_n. \quad (24)$$

We first note that a given trajectory can be used to evaluate all terms with order lower than n. Trajectories, which leave the device domain after k iterations, give zero contribution to the sample mean (D2) for any term with $n > k$. Nevertheless, such trajectories are counted as independent realizations in the denominator N of the sample mean. It follows that a given trajectory can be used for evaluation of all terms in (18): a trajectory that begins at a domain boundary and ends at a domain boundary becomes an independent realization of the random variable $\psi_{\langle A\rangle}$ in (24). The sample mean over N such trajectories estimates $\langle A\rangle$.

IV. THE PARTICLE MODEL

A. Classical transport and ergodicity

In the case of Boltzmann transport, the Wigner potential completes the Liouville operator with a force term. The trajectories have the general form (A1). On the right-hand side of (2), only the two terms related to the phonons remain. This allows one to formally set $\gamma = 0$, $\mu = \lambda$ in (7). The right-hand side of (22) reduces to p_{ph}. It is easily seen that the numerical trajectories coincide with the real trajectories of the Boltzmann carriers evolving in the device. Indeed, the generation of the initial point of the trajectory corresponds to an injection of a classical particle. p_t becomes the usual classical probability for the free-flight duration due to phonons. The scattering is determined by the phonon scattering rate S through p_{ph} in (22). The weight W_n in (24) remains unity for all n. The domain indicator takes into account only the trajectories that are inside the device. Thus, numerical particles contribute to the averages in the same way as Boltzmann carriers in the Single-Particle MC method. We conclude that the resulting particle picture coincides with the picture of classical particles that is emulated by the the device MC method.

An important property follows from this equivalence. According to the classical transport theory, the distribution function f is given by the relative number n_Ω/N_D of the particles in any given unit phase-space volume Ω with indicator θ_Ω. Here N_D is the number of all carriers, so that f is normalized to unity in the device. This is, in fact, the basis of obtaining the physical averages by the Ensemble Monte Carlo method. Under conditions of stationary transport, and by assuming ergodicity, the ensemble average is replaced by a time average over a single particle: f is given by the relative time spent by the particle in Ω. This is the basis for obtaining averages by the Single Particle MC method. We show that this result follows from the present approach. Indeed, N independent trajectories can be regarded as having been obtained by N consecutive reinjections of a single particle that evolves in the device until exiting. By setting $A = \theta_1$, from (23), (24), and (D2), it follows that $n_\Omega = (\Phi/N)t_\Omega$, where t_Ω is the total time spent by the particle in Ω. Accordingly, $N_D = (\Phi/N)T_D$, where T_D is the total time spent by the particle in the device. Then the distribution function $f = n_\Omega/N_D = t_\Omega/T_D$ is estimated by the relative time spent in Ω. We conclude that the ergodicity is not required to be assumed, as it follows from the stationary conditions of the transport.

B. Coherent transport

The coherent transport, which considers only events of quantum interaction, is obtained by setting $\lambda = 0$, $\mu = \gamma$ in (7). According to the term in brackets in (22), these are scattering events that change the statistical weight in (24). One can estimate the mean accumulated weight W from the mean time T a trajectory spends in the device. T is given by the sum of all free-flight times. The number n of the scattering events is then $n = T\gamma$, and W is estimated as

$$W = \pm(3)^n = \left(1 + \frac{2\gamma}{\gamma}\right)^n = \left(1 + \frac{2\gamma T}{n}\right)^n \simeq \exp(2\gamma T).$$

It follows that the mean weight, and thus the variance, grows exponentially with the magnitude of the Wigner potential and the dwell time T. This result is in accordance with the exponential growth with time of the variance of the MC approach to Feynman path integrals.[79] If the device dimen-

sions are larger than ten nanometers, T is commonly larger than a picosecond, while $\gamma \simeq 10^{15}$ s^{-1} for a 0.3 eV potential barrier. This precludes the application of the approach to mesoscopic devices. This version of the approach has been applied to tunneling through one-order-of-magnitude-smaller potential barriers.[80]

An alternative interpretation is needed in order to solve the problem with the accumulated weight. We explore the idea of particle splitting, which is an established approach,[81] for statistical enhancement in classical Monte Carlo simulations. A particle entering a rarely visited region of the phase-space can be split into subparticles, each of which carries a fraction of the particle weight. This can be achieved by setting $p_2=1$ in (22). This modification changes entirely the interpretation of the quantum term. Now, the state $Q'=(\mathbf{r}',\mathbf{k}')$, which enters the interaction, gives rise to three states, so that the Wigner potential is understood as a generational term. After the interaction, the initial particle survives in the same state, due to the delta function in the brackets of (22). Two additional particles are generated by $p_w^{\pm}$ in states $Q^{\pm}=(\mathbf{r}',\mathbf{k}^{\pm})$.[82] The trajectory now branches so that the weight carried by any branch keeps a constant magnitude and can change only the sign. It can be seen from (19) that the branching corresponds to a splitting of the integral into three integrals. Hence each branch continues with a free flight to contribute to the sample mean of ψ_1. One of the contributions carries the minus sign of p_w^-. It is beneficial to assign a sign to the particle associated to each trajectory. Then the following transport process can be imagined. A positive particle is injected from the device boundaries. It drifts over a trajectory (4) until the interaction time generated according p_t is reached. The particle does not "feel" the Wigner potential, because after the interaction it remains in the same state. The next drift process continues on the same trajectory. The action of the potential is realized through a creation of two new particles in two phase-space states. The particle related to p_w^+ (p_w^-) has the same (the opposite) sign as the primary particle. The created particles follow the same evolution process over their own trajectories.

After each individual interaction, any positive (negative) particle contributes to the estimator of $\psi_{\langle A\rangle}$ with $+(-)\Phi\psi_A$. Two particles that are in the same phase-space point follow a common trajectory. If they have opposite signs, they give opposite contributions to $\psi_{\langle A\rangle}$. Moreover such particles create with the same probability for any point of the phase-space particles with the opposite sign. The net contribution of such particles to the physical averages as well as to the generation process is zero. It follows that particles with the opposite sign that meet in the phasespace can be annihilated. The coherent transport is characterized by processes of generation and annihilation of positive and negative particles.

C. Quantum transport with dissipation

The above two limiting cases of the Wigner-Boltzmann transport can be combined without interference into a general picture of quantum transport with dissipation. The phonon interaction is inserted on top of the coherent picture and affects the dynamics of the particles. They no longer remain on a single trajectory throughout the device, but are scattered to different trajectory pieces after each process of drift. According to (22), the events of phonon and quantum interactions are complementary. The action of the Wigner potential on the interacting particle is equivalent to a self-scattering event, as it does not change the trajectory. From this analogy, it follows that the duration of the free flight on a given trajectory depends only on λ. The after-scattering state is selected by the phonon rate S through p_{ph}. We conclude that particles have the same Boltzmann-like behavior in both classical and quantum regions. The quantum character of the transport is marked by the generation process and the sign of the particles.

The possibility to annihilate particles with opposite sign also remains true. The reason is that the evolution does not change its Markovian character in the case of phonons. Particles at a given phasespace point still have a common probabilistic future, and the considerations from the previous section apply. A condition for this is that the phonons are treated in a classical way. If the interaction is quantum, this property is not generally true because of the memory character of the evolution.

We note that other interpretations conserving the absolute weight on a trajectory are possible. The interaction with the Wigner potential can be chosen twice as rare on the expense that four particles are created per such event. A reformulation of (22) can lead to events where quantum and phonon interactions occur in the same instances.[83] The proposed particle picture is the most straightforward one, which follows from this approach.

V. NUMERICAL ASPECTS

The practical application of the above model is not straightforward. The trajectories can be simulated only sequentially so that the generated particles must be stored for further processing. This is a step common with the classical split algorithm. In the latter, the stored subparticles are removed in a subsequent simulation step and reinjected from the boundaries. This is not possible with quantum particles. A single particle injected from the boundaries creates, during the evolution throughout the device, a first generation of stored particles. The subsequent removal of the stored particles gives rise to a second generation, etc. The process can be infinitely continued; the steps of injection and removal cannot be deduced from Heuristic considerations.

The solution is given by an algorithm based on decomposition of $\tilde{K}=L+M$ into two operators L and M,

$$L = p_t\theta_D(p_\lambda p_{ph} + p_\gamma p_\delta), \quad M = p_t\theta_D p_\gamma (p_w^+ - p_w^-). \quad (25)$$

In L, we recognize the Boltzmann evolution operator accounting for a phonon interaction augmented by a self-scattering term. The self-scattering process does not affect the physical picture; it can be switched on and off without affecting the shape of the trajectories. Usually the process is used in the device MC method for computational convenience. Thus L gives rise to a classical evolution of the trajectories.

The above operators obey the equality[84]

$$(I-\tilde{K})^{-1}=(I-L)^{-1}+(I-\tilde{K})^{-1}M(I-L)^{-1}. \quad (26)$$

Equation (26) is proved by multiplication by $(I-L)$ from the left and $(I-\tilde{K})$ from the right. A substitution of (26) into (18), and regrouping the terms of the series leads to

$$\langle A\rangle=(\mathbf{v}_\perp f_b,(I-\tilde{K})^{-1}\tilde{A})=\sum_{k=0}^{\infty}(\mathbf{v}_\perp f_b,((I-L)^{-1}M)^k(I-L)^{-1}\tilde{A}). \quad (27)$$

We now derive an algorithm that evaluates the consecutive terms in (27). The first term $(A)_1=[\mathbf{v}_\perp f_b,(I-L)^{-1}A]$ can be evaluated by classical means; N_1 particles are consecutively injected from the boundaries and simulated throughout the device. The sample mean for $(A)_1$ can be formulated with the help of (D2) and (24)

$$(A)_1\simeq\frac{\Phi}{N_1}\sum_{j,n}\prod_n\theta_{Dn}^{(j)}\psi_{An}^{(j)}. \quad (28)$$

The second term $(A)_2=[\mathbf{v}_\perp f_b,(I-L)^{-1}M(I-L)^{-1}\tilde{A}]$ can be reformulated by introducing a delta function

$$(A)_2=[P_1,(I-L)^{-1}\tilde{A}],$$

$$P_1(Q)=[\mathbf{v}_\perp f_b,(I-L)^{-1}M\delta](Q). \quad (29)$$

We show that P_1, up to a prefactor, has the meaning of density of stored particles. The series $(I-L)^{-1}M\delta$ is obtained by a replacement of $\tilde{A}=p_t\psi_A$ by $M\delta$ in the consecutive terms of $(I-L)^{-1}\tilde{A}$. This is actually equivalent to replacing ψ_A by $p_\lambda(p_w^+-p_w^-)\delta$, which corresponds to a generation with probability p_λ of two particles in states Q^+ and Q^-. The delta function at the end gives rise to a projection to the phase-space point Q, which is achieved by the random variable $\psi_\delta=\delta(Q-Q^+)-\delta(Q-Q^-)$. We first note that both $(A)_1$ and P_1 can be sampled over common classical trajectories constructed with the help of L. The generation of the particles for P_1 occurs at the same instances with the self-scattering events. Second, $\psi_\delta dQ$ is 1, (-1) if a particle is generated into a point Q^+, (Q^-) that belongs to the domain dQ around Q. Since $Q^+\neq Q^-$, the case where both points belong to dQ can be avoided by the limit $dQ\to 0$. The estimator of P_1 is (28) with ψ_δ in the place of ψ_A. It follows that P_1 is the density w_p of stored particles multiplied by the prefactor Φ/N_1. We note that the annihilation of positive and negative particles in a given Q is formally proven by this approach.

The average $(A)_2$ in (29) differs from $(A)_1$ only by the boundary term, which is now replaced by P_1. It can be evaluated with minor modifications of the algorithm: the trajectories should now begin from the device volume. The estimator depends on the choice of the density p for selection of the initial trajectory points Q:

$$(A)_2\simeq\frac{1}{N_2}\sum_{j,n}\frac{P_1(Q^{(j)})}{p(Q^{(j)})}\prod_n\theta_{Dn}^{(j)}\psi_{An}^{(j)}. \quad (30)$$

Here, N_2 is the number of trajectories used. A convenient choice for p is

$$p(Q)=|P_1(Q)|/\|P_1\|,\quad \|P_1\|=\frac{\Phi}{N_1}\int|w_p(Q)|dQ.$$

If N_2 is chosen to be equal to the number of all stored particles given by the last integral above, (30) becomes

$$(A)_2\simeq\frac{\Phi}{N_1}\sum_{j,n}\mathrm{sign}P_1(Q^{(j)})\prod_n\theta_{Dn}^{(j)}\psi_{An}^{(j)}. \quad (31)$$

The algorithm discharges the device from the stored particles; the number of trajectories that initiate from dQ and their initial sign corresponds to the number and sign of the stored particles inside. The value of each contribution to the sum in (31) is calculated in the classical way and multiplied by the sign of the corresponding particle. At this step, we can evaluate the function

$$P_2(Q)=[P_1,(I-L)^{-1}M\delta](Q). \quad (32)$$

By repeating the arguments used for P_1 it is seen that P_2 is the density of the secondary stored particles multiplied by Φ/N_1. The third term in (27) is then expressed as an inner product with P_2: $(A)_3=[P_2,(I-L)^{-1}A]$. The step used for $(A)_2$ can be repeated for $(A)_3$, and so forth. An iterative algorithm is obtained, which computes the consecutive terms of the series (27) by an initial injection of N_1 particles from the boundary and consecutive steps of storing and removing particles from the device. Since the estimators of the consecutive terms in the series have to be summed at the end, only a single estimator is necessary for the evaluation of $\langle A\rangle$. The prefactor $C=\Phi/N_1$ appears in the estimators of all terms and can be determined in the manner discussed in Sec. IV A.

The number N_1 of the boundary particles must be chosen sufficiently large in order to attain a reliable approximation of the first term in (27). The evaluation of all remaining terms depends on this choice. In order to be less dependent on the initial guess of this value, the algorithm can be modified as follows. A moderate value of N_1 is chosen. The steps of injection from the boundaries alternate with the steps of discharge of the device. Particles stored from a boundary injection are added to the particles stored from the previous injection in the device. This process of accumulation of the particles in the device follows the scheme

$$b\to w_{p1}\underset{\to}{\overset{b\to}{}}(w_{p1}+w_{p2})\underset{\to}{\overset{b\to}{}}(w_{p1}+w_{p2}+w_{p3})\ldots,$$

where $b\to$ means boundary injection, which gives rise to w_{p1} in all brackets. The rest of the terms w_{pi} in each bracket is obtained from the previous step of discharging the device, which is denoted by the bottom arrow. After the Rth boundary injection, the number of stored particles in dQ around Q is approximately $\Sigma_{k=1}^{R}w_{pk}(Q)dQ$, which can be proven by induction. The subsequent step of injection from the device completes the estimation of the functional $\Sigma_{k=1}^{R+1}(A)_k$. We take the mean of these estimates at steps $0,1,\ldots,R$ as an approximation of the functional $\langle A\rangle$. Since the functional $(A)_k$ is summed $R-k$ times and the ratio $(R-k)/R$ tends to 1 as R tends to infinity, our estimate of $\langle A\rangle$ is asymptotically correct. The modified algorithm assists the process of annihila-

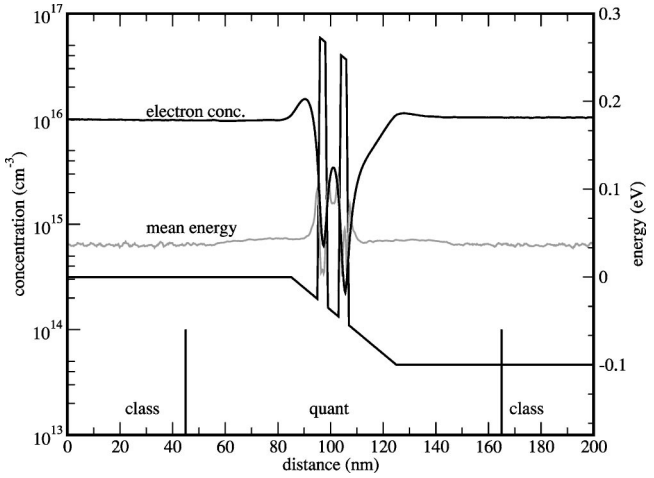

FIG. 2. Electron concentration and mean electron energy at T = 300 K and 0.1 V bias.

tion and suppresses the development of rare events that can degrade the statistics.

VI. SIMULATION RESULTS AND DISCUSSIONS

The proposed particle method has been investigated by simulations of benchmark resonant tunneling diodes (RTD). The features and relevance of the method with respect to processes of tunneling and dissipation is studied. The method has been successively compared to the comprehensive solver NEMO-1D based on nonequilibrium Green's functions.[85] A common physical model is applied. The carrier transport in the RTDs is cosidered one-dimensional; the phase space is defined by the single device coordinate and the complete three-dimensional space of wave vectors. The numerical counterpart of the definition (3) is given by a discrete Fourier transform defined in the interval $(-L_c/2, L_c/2)$ where L_c is called coherence length. The standard analytical band scattering model used for the simulated RTDs includes polar optical and acoustic deformation potential scattering, assuming parameter values for GaAs. The position-independent effective mass is $0,067m_0$.

Device I, which is shown in Fig. 2, has been taken from Shifren *et al.*[62] The two barriers are 0.3 eV high and 3 nm thick; the well width is 5 nm. A linear voltage drop is assumed in a region from 10 nm before the emitter barrier to 19 nm after the collector barrier (a total distance of 40 nm). In the contact regions, the doping is 10^{16} cm^{-3}. The coherence length used in the simulations is $L_c = 80$ nm. The Wigner generation rate γ reaches fairly high values in this structure (Fig. 3) and is used to distinguish between classical and quantum regions. Comparing this rate with the much smaller phonon scattering rate is a quantitative measure of the fact that quantum-interference effects are dominant.

The electron concentration and the mean energy are shown in Fig. 2. These are nearly constant in the contact regions. A smooth transition between classical and quantum regions is demonstrated despite the strong onset of γ, shown in Fig. 3 for two values of the applied bias.

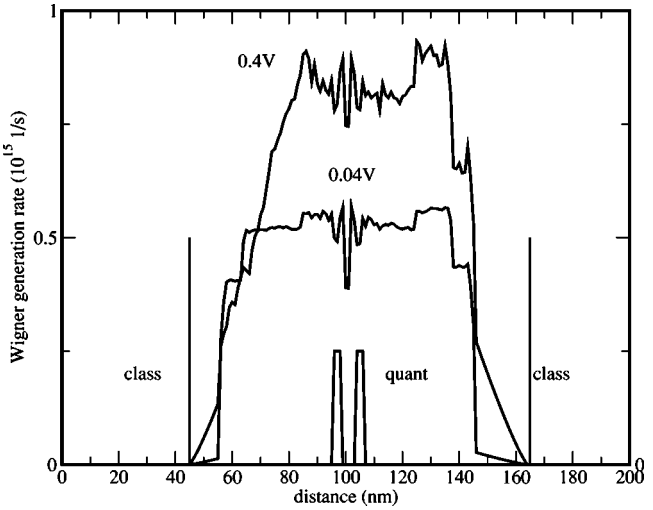

FIG. 3. Pair generation rate $\gamma(x)$ caused by the Wigner potential.

Phonon scattering strongly affects the current-voltage characteristics, as can be seen in Fig. 4. In comparison to the coherent case, the scattering leads to an increase of the valley current and a shift of the resonance voltage. This effect is due to a repopulation of the electron states in the emitter. Inelastic scattering events dissipate the energy of the electrons entering from the left-hand reservoir. Propagating electrons fall into the lower energy states in the potential notch on the left-hand side of the barrier (Fig. 2) and contribute to the current. The large difference in the valley current can be explained with the electron concentration in off-resonance conditions (Fig. 5). The scattering with phonons leads to significantly higher concentration in the notch. A quasibound state is formed, and the injection in the double barrier is increased. In resonance conditions, where the applied voltage is lower, such a bound state does not form and very similar electron concentrations are observed for the coherent and noncoherent case, which is shown in Fig. 6.

These effects cannot be accounted for by elementary tunneling theories, which assume the notch states to be in equi-

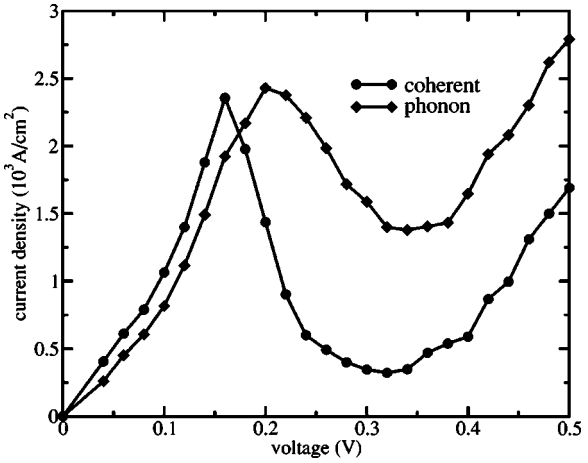

FIG. 4. Influence of phonon scattering on the I/V characteristics of the RTD.

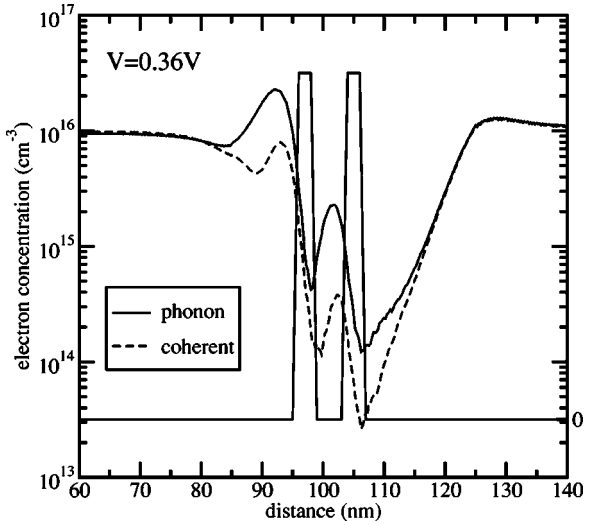

FIG. 5. Electron concentration in off-resonance conditions.

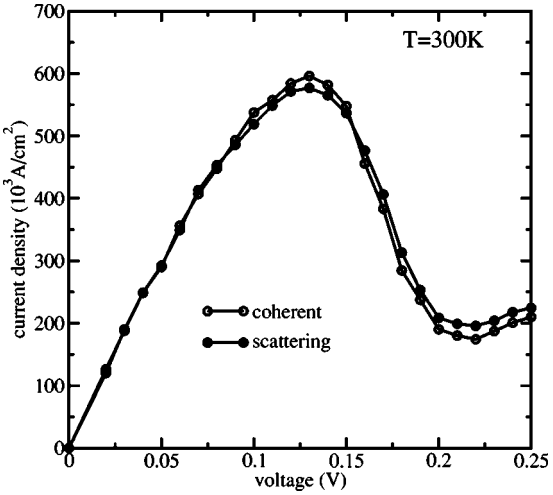

FIG. 7. Phonon scattering has a small effect on the device characteristics.

librium with the right-hand reservoir.[9] From these considerations, it follows that the effect of phonons should be negligible in a device having a flat potential on the left-hand side of the barrier. Tsuchiya *et al.*[86] have investigated such an RTD, assuming a linear potential drop only across the central device. The two 2.825 nm barriers are 0.27 eV high, and the well width is 4.52 nm. The doping level in the contacts is 2×10^{18} cm^{-3}. The I/V characteristics of Device II with and without phonon scattering are shown on Fig. 7. Consistently, phonons cause only a small effect on the device behavior.

The I/V coherent characteristics of Device II, as obtained by the particle method (QMC) and NEMO, are shown in Fig. 8. Compared to NEMO, a slightly higher peak current is obtained by the particle method. The results show a good qualitative agreement.

The coherence length L_c has to be selected carefully in Wigner transport simulations. A comparative study, the results of which are plotted in Fig. 9, shows that only a sufficiently large value gives realistic results. The finite difference (FD) result is taken from Tsuchiya.[86] A too small value for L_c leads to an overestimation of the valley current. The requirement that the coherence length must be chosen sufficiently large can be linked to the completeness relation of the discrete Fourier transform: $\Delta k_x = \pi/L_c$. A small L_c results in a coarse grid in the momentum space so that resonance peaks are likely not resolved properly. A large L_c can create problems for the finite difference methods. The sparsity pattern of the matrix related to the Wigner potential is very poor. An increase of the matrix dimension, evaluated by $L_c/\Delta x$, can lead to prohibitive memory requirements. The stochastic approach does not require operations with matrices, so that the number of points can be easily increased. In this work, the Wigner potential is discretized using approximately 10^3 points. Moreover, the particle method accounts for dissipation processes at a rigorous kinetic level. As compared to tools based on tunneling theories, or a coherent Wigner function, this method avoids the usual relaxation time approximation.

Figure 10 shows the carrier distribution for different bias points in a third type of RTD. The 4 nm well of Device III is

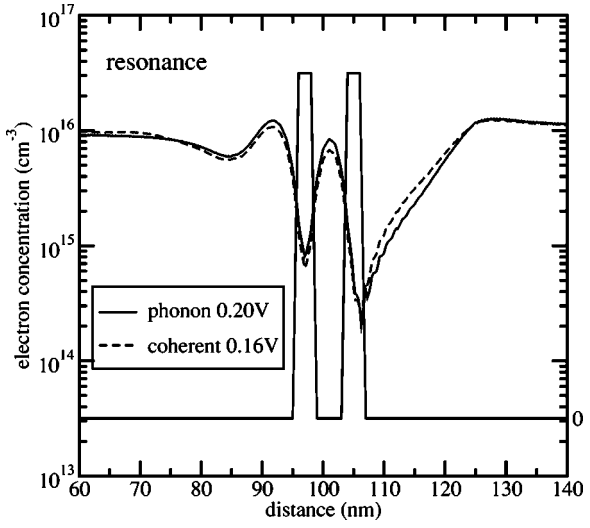

FIG. 6. Electron concentration in resonance conditions.

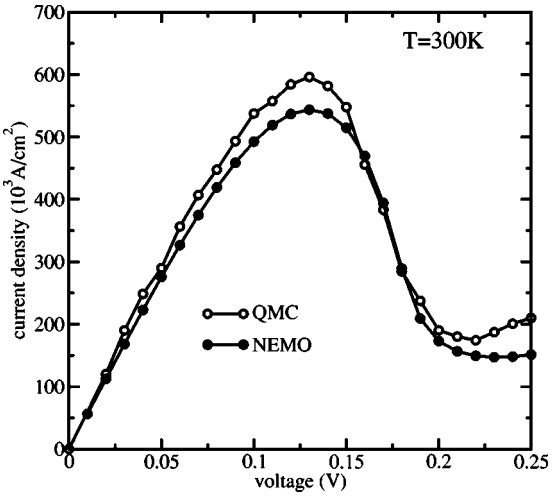

FIG. 8. Coherent I/V characteristics obtained by the particle method (QMC) and NEMO.

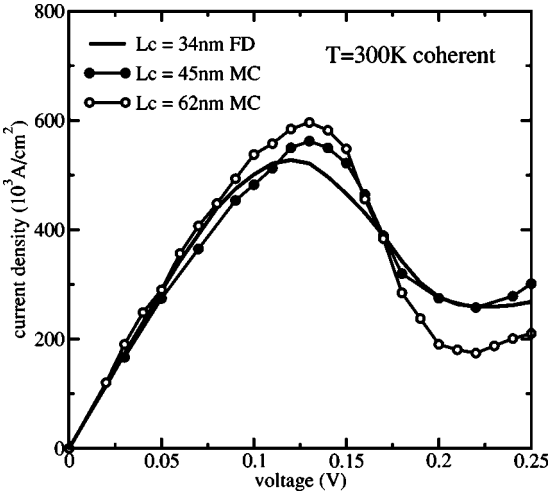

FIG. 9. Effect of the coherence length in Wigner simulations.

enclosed by 3 nm barriers with 0.49 eV height. The relevant simulation domain is 300 nm, the quantum domain of 120 nm corresponds to $L_c=60$ nm. 1200 k_x points are used and the spacing in the x direction is $\Delta x=0.5$ nm. At the resonance voltage of 1.2 V, the concentration in the quantum well is considerably higher than in the off-resonance condition at 1.6 V. The concentration in the depletion region left of the barriers depend on the injected current from the right and thus correlates with the concentration in the well. The 1.6 V curve provides an instructive example illustrating the numerical noise in the region where the carrier density drops more than three orders of magnitude.

The numerical noise of the stochastic approach poses a natural limit for the resolution of the simulated averages. Physical conditions, where physical quantities vary over several orders of magnitude, are unattainable by this method. High-performance devices with a high peak/valley ratio are beyond the scope of the particle approach. Therefore, it is concluded that the present method can bridge the gap between classical device simulations and pure quantum-ballistic simulation. This method provides a relevant description of the stationary carrier transport in nanostructures at moderate physical conditions. Effects of tunneling and dissipation are incorporated with equal numerical accuracy and the results agree with other rigorous simulation tools.

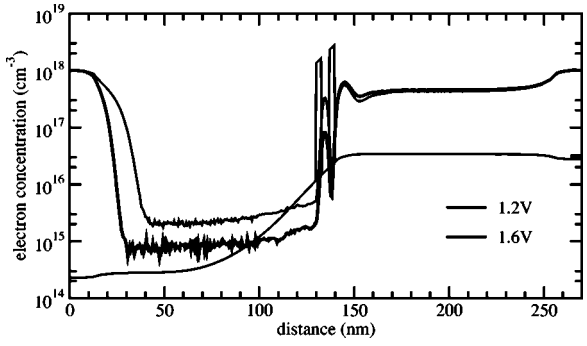

FIG. 10. Electron concentration profiles in Device III.

VII. CONCLUSIONS

A particle method has been associated with Wigner-Boltzmann transport in small semiconductor devices. This method accounts for dissipation and interference phenomena by two alternative processes of transport of quasiparticles. Dissipation caused by interaction with phonons appears as consequence of both drift and scattering processes corresponding to semiclassical Boltzmann transport. Interference effects due to the Wigner potential are associated with generation of couples of particles having statistical weight ±1. In this picture, the Wigner equation with a Boltzmann scattering term has been interpreted as a Boltzmann equation augmented by a generation term.

For sufficiently smooth potentials, the Wigner operator can be approximated by a classical force term, which is included in the Liouville operator. The present particle method simplifies gradually to the classical MC method when the classical limit is approached.[87] A seamless transition between classical and quantum regions results. It has been shown that the ergodicity of the classical transport process follows from the assumed stationary transport conditions.

In the regions of quantum transport, the challenge of employing the method is to handle the avalanche of the particles properly. The problem has been solved for stationary conditions: Particles of opposite sign and a sufficiently small distance in phase space annihilate one another in the course of a simulation. Positive or negative particles can accumulate in phase space domains. The accumulation of negative particles is in accordance with the negative values which characterize the Wigner function especially in the regions of strong potential variations. The approach is suitable for moderate physical conditions of mixed coherent and dissipative transport.

APPENDIX A

The freedom of the choice of the initialization time directly follows from the fact that the wave vector $\mathbf{k}$ remains constant in time for an unaccelerated trajectory. Actually this property is valid for a general Liouville operator with a force term that does not depend explicitely on time. A useful relation, which follows from the stationarity assumed, will be derived.

A trajectory is initialized by a phase space point $(\mathbf{k},\mathbf{r})$ and a time t_0

$$\mathbf{R}(t;t_0,\mathbf{r},\mathbf{k}) = \mathbf{r} + \int_{t_0}^{t} \mathbf{v}(\mathbf{K}(y;\cdot))dy,$$

$$\mathbf{K}(t;t_0,\mathbf{r},\mathbf{k}) = \mathbf{k} + \int_{t_0}^{t} \mathbf{F}(\mathbf{R}(y;\cdot))dy. \qquad \text{(A1)}$$

The order of t_0 and t is irrelevant. A trajectory is called forward if the evolution time is greater than the initialization time $t>t_0$. Otherwise the trajectory is called backward. The limits of the time integration in (A1) can be exchanged by changing the sign of the integrals. To describe a time-invariant system an absolute time scale is not needed. Only

the time difference between two consecutive events is important. Invariance under time translation can be proven

$$\mathbf{R}(t+\tau;t_0+\tau,\mathbf{r},\mathbf{k})=\mathbf{R}(t;t_0,\mathbf{r},\mathbf{k}),$$

$$\mathbf{K}(t+\tau;t_0+\tau,\mathbf{r},\mathbf{k})=\mathbf{K}(t;t_0,\mathbf{r},\mathbf{k}). \tag{A2}$$

This property will be used to conveniently adjust the time reference $t_0=0$ for each trajectory. A shortcut notation $\mathbf{r}(t)=\mathbf{R}(t;0,\mathbf{r},\mathbf{k})$ and $\mathbf{k}(t)=K(t;0,\mathbf{r},\mathbf{k})$ is introduced. A particularly useful relation is obtained from (A2)

$$\int d\mathbf{r}d\mathbf{k}\int_{-T}^{0}dt\phi(\mathbf{r},\mathbf{k},\mathbf{r}(t),\mathbf{k}(t))$$

$$=\int d\mathbf{r}'d\mathbf{k}'\int_{0}^{T}dt\phi(\mathbf{r}'(t),\mathbf{k}'(t),\mathbf{r}',\mathbf{k}'). \tag{A3}$$

Here $\mathbf{r}(t)$, $\mathbf{k}(t)$ is a backward trajectory initialized by $\mathbf{r},\mathbf{k}$, while $\mathbf{r}'(t),\mathbf{k}'(t)$ is a forward trajectory initialized by $\mathbf{r}',\mathbf{k}'$. The relation is proven by introducing new integration variables $\mathbf{r}'=\mathbf{R}(t;0,\mathbf{r},\mathbf{k})$, $\mathbf{k}'=\mathbf{K}(t;0,\mathbf{r},\mathbf{k})$. Then

$$\mathbf{r}=\mathbf{R}(0;t,\mathbf{r}',\mathbf{k}')=\mathbf{R}(-t;0,\mathbf{r}',\mathbf{k}'),$$

$$\mathbf{k}=\mathbf{K}(0;t,\mathbf{r}',\mathbf{k}')=\mathbf{K}(-t;0,\mathbf{r}',\mathbf{k}').$$

According to the Liouville theorem the phase volume is invariant under this transformation: $d\mathbf{r}'d\mathbf{k}'=d\mathbf{r}d\mathbf{k}$. The last step is to reverse the time by switching the sign of t and to recall that $\mathbf{R}(t;0,\mathbf{r}',\mathbf{k}')$ and $\mathbf{K}(t;0,\mathbf{r}',\mathbf{k}')$ are forward trajectories initialized by $(\mathbf{r}',\mathbf{k}')$ and denoted by $\mathbf{r}'(t),\mathbf{k}'(t)$. In particular, for a given function μ it follows that

$$\int_{-T}^{0}\mu(\mathbf{r}(y),\mathbf{k}(y))dy=\int_{0}^{T}\mu(\mathbf{r}'(y),\mathbf{k}'(y))dy. \tag{A4}$$

APPENDIX B

To obtain the integral form of the stationary Wigner-Boltzmann equation, we consider a given phase space point $(\mathbf{r},\mathbf{k})$. This point determines uniquely a phase space trajectory, $(\mathbf{r}(t),\mathbf{k}(t))$ in backward parametrization, $t\leqslant 0$. Consider the parametrized Eq. (5):

$$[\mathbf{v}(\mathbf{k}(t))\cdot\nabla_{\mathbf{r}(t)}+\mu(\mathbf{r}(t),\mathbf{k}(t))]f_w(\mathbf{r}(t),\mathbf{k}(t))$$

$$=\int d\mathbf{k}'\Gamma(\mathbf{r}(t),\mathbf{k}(t),\mathbf{k}')f_w(\mathbf{r}(t),\mathbf{k}'). \tag{B1}$$

If both sides of (B1) are multiplied by the integrating factor $\exp[-\int_t^0\mu(\mathbf{k}(y),\mathbf{r}(y))dy]$, the left-hand side represents a total time derivative

$$\frac{d}{dt}\exp\left(-\int_t^0\hat{\mu}(y)dy\right)\hat{f}(t)=\exp\left(-\int_t^0\hat{\mu}(y)dy\right)\hat{\Gamma}[f](t). \tag{B2}$$

Here $\hat{\Gamma}[f](t)$ denotes the right-hand side of (B1) and $\hat{\mu}(y)=\mu(\mathbf{k}(y),\mathbf{r}(y))$, $\hat{f}(t)=f_w(\mathbf{r}(t),\mathbf{k}(t))$. This equation can be integrated straightforwardly. The upper bound of integration should be $t=0$ to obtain $\hat{f}(0)=f(\mathbf{k},\mathbf{r})$, the value of f at the given phase space point. The lower-time bound has to be chosen such that the functions $\mathbf{K}(t)$ and $\mathbf{R}(t)$ take on values at which the distribution function is known. In the steady state the distribution function is known only at the device boundary. An appropriate lower time bound is therefore the time, say t_b^-, at which the trajectory crosses the simulation domain boundary. Apparently, this time depends on the point $\mathbf{k}$, $\mathbf{r}$ under consideration. For trajectories closed in D the time t_b is $-\infty$. Integration of (B2) in the time bounds discussed above results in the integral form of the stationary Wigner-Boltzmann equation (8).

APPENDIX C

We first introduce the function $B(\mathbf{r})=0$, which defines the domain boundary. This gives an implicit definition of the boundary time as a root of $B(\mathbf{r}(t'))=0$. The equality

$$\delta(t'-t_b)=\delta[B(\mathbf{r}(t'))]|\nabla_{\mathbf{r}}B(\mathbf{r}_b)||\mathbf{v}_\perp(\mathbf{k}(t_b))| \tag{C1}$$

follows directly from the properties of the δ function. The normal velocity component $\mathbf{v}_\perp$ appears since the gradient of B is normal to the domain boundary in the crossing point with the trajectory. Using the right-hand side of (C1) in the augmented equation (16) gives

$$\langle A\rangle=\int_{-\infty}^{0}dt'\int_D d\mathbf{r}\int d\mathbf{k}\delta[B(\mathbf{r}(t'))]|\nabla_{\mathbf{r}}B(\mathbf{r}_b)||\mathbf{v}_\perp(\mathbf{k}(t_b))|$$

$$\times f_b(\mathbf{r}(t_b^-),\mathbf{k}(t_b^-))\exp\left(-\int_{t_b^-}^{0}\mu(\mathbf{r}(y),\mathbf{k}(y))dy\right)g(\mathbf{k},\mathbf{r}).$$

The variables $(\mathbf{k}',\mathbf{r}')$ are changed to $(\mathbf{k}_b=\mathbf{k}(t'),\mathbf{r}''=\mathbf{r}(t')$ and (A3) is applied. The $\mathbf{r}''$ integral is accounted with the aid of the delta function

$$\int_D\delta(B(\mathbf{r}''))\phi(\mathbf{r}'')d\mathbf{r}''=\oint\frac{\phi(\mathbf{r}_b)}{|\nabla_{\mathbf{r}}B(\mathbf{r}_b)|}d\sigma(\mathbf{r}_b),$$

where ϕ is a test function.[88] This accomplishes the change from volume to boundary integration leading to (17).

APPENDIX D

Consider a random variable ψ which takes values $\psi(y)$ with probability density $p_\psi(y)$. Here y is a multidimensional point. The expectation value E_ψ of ψ is given by

$$E_\psi=\int dyp_\psi(y)\psi(y). \tag{D1}$$

The simplest Monte Carlo method evaluates E_ψ by performing N independent realizations of the probability density p_ψ. Generated are N points $y_1,\ldots,y_N$, called sampling points for the random variable ψ. The sample mean η estimates the expectation value E_ψ

$$E_\psi \simeq \eta = \frac{1}{N}\sum_{i=1}^{N} \psi(y_i) \tag{D2}$$

with a precision that depends on the number of independent realizations N and the variance of σ_ψ of the random variable. According to the "rule of the three sigma":[89]

$$P\left\{ |E_\psi - \eta| \leqslant \frac{3\sigma_\psi}{\sqrt{N}} \right\} \simeq 0.997 \tag{D3}$$

the probability for η to be inside the interval $3\sigma_\psi/\sqrt{N}$ around E_ψ is very high (0.997).

The concept of the Monte Carlo approach for evaluation integrals is to present a given integral as an expectation value

$$I = \int f(y)dy = \int p(y)\frac{f(y)}{p(y)}, \quad p(y) \geqslant 0, \quad \int p(y)dy = 1 \tag{D4}$$

of the random variable $\psi = f/p$. The probability density function p can be arbitrary, but admissible for f: $p \neq 0$ if $f \neq 0$. Different random variables can be introduced, depending on the choice of p. All of them have the same expectation value I but different variance and higher moments. It can be shown that the lowest variance is obtained if p is chosen proportional to $|f|$.

[1] E. Wigner, Phys. Rev. **40**, 749 (1932).
[2] J. Rammer, Rev. Mod. Phys. **63**, 781 (1991).
[3] B. Biegel and J. Plummer, Phys. Rev. B **54**, 8070 (1996).
[4] V. I. Tatarskii, Sov. Phys. Usp. **26**, 311 (1983).
[5] U. Ravaiolli, M. A. Osman, W. Poetz, N. C. Kluksdahl, and D. K. Ferry, Physica B & C **134B**, 36 (1985).
[6] W. Frensley, Phys. Rev. B **36**, 1570 (1987).
[7] N. C. Kluksdahl, W. Poetz, U. Ravaiolli, and D. K. Ferry, Superlattices Microstruct. **3**, 41 (1987).
[8] N. C. Kluksdahl, A. M. Kriman, D. K. Ferry, and C. Ringhofer, Phys. Rev. B **39**, 7720 (1989).
[9] W. Frensley, Rev. Mod. Phys. **62**, 745 (1990).
[10] K. Gullapalli, D. Miller, and D. Neikirk, Phys. Rev. B **49**, 2622 (1994).
[11] F. A. Buot and K. L. Jensen, Phys. Rev. B **42**, 9429 (1990).
[12] F. Rossi, C. Jacoboni, and M. Nedjalkov, Semicond. Sci. Technol. **9**, 934 (1994).
[13] P. Bordone, M. Pascoli, R. Brunetti, A. Bertoni, and C. Jacoboni, Phys. Rev. B **59**, 3060 (1999).
[14] M. Nedjalkov, R. Kosik, H. Kosina, and S. Selberherr, Microelectron. Eng. **63**, 199 (2002).
[15] I. Levinson, Sov. Phys. JETP **30**, 362 (1970).
[16] J. R. Barker and D. K. Ferry, Phys. Rev. Lett. **42**, 1779 (1979).
[17] M. Nedjalkov and I. Dimov, Math. Comput. Simul. **47**, 383 (1998).
[18] J. Schilp, T. Kuhn, and G. Mahler, Phys. Rev. B **50**, 5435 (1994).
[19] T. Gurov and P. Whitlock, Math. Comput. Simul. **60**, 85 (2002).
[20] M. Nedjalkov, H. Kosina, S. Selberherr, and I. Dimov, VLSI Des. **13**, 405 (2001).
[21] T. V. Gurov, M. Nedjalkov, P. A. Whitlock, H. Kosina, and S. Selberherr, Physica B **314**, 301 (2002).
[22] M. Nedjalkov, R. Kosik, H. Kosina, and S. Selberherr, *Proceedings of 2001 First IEEE Conference on Nanotechnology* (Hawaii, IEEE New York, 2001), Vol. 277.
[23] J. Barker, Solid-State Electron. **21**, 267 (1978).
[24] J. Barker and D. Ferry, Solid-State Electron. **23**, 519 (1980).
[25] J. Barker and D. Ferry, Solid-State Electron. **23**, 531 (1980).
[26] D. Ferry and J. Barker, Solid-State Electron. **23**, 545 (1980).
[27] C. Jacoboni, A. Bertoni, P. Bordone, and R. Brunetti, Math. Comput. Simul. **55**, 67 (2001).
[28] A. Bertone, P. Bordone, R. Brunetti, C. Jocoboni, and N. Sano, Physica B **272**, 299 (1999).
[29] R. Brunetti, C. Jacoboni, and F. Rossi, Phys. Rev. B **39**, 10 781 (1989).
[30] P. Bordone, A. Bertoni, R. Brunetti, and C. Jacoboni, Math. Comput. Simul. **62**, 307 (2003).
[31] W. Porod and D. Ferry, Physica B & C **134B**, 137 (1985).
[32] M. Nedjalkov, H. Kosina, S. Selberherr, and I. Dimov, VLSI Des. **13**, 405 (2001).
[33] P. Lipavski, F. Khan, F. Abdolsalami, and J. Wilkins, Phys. Rev. B **43**, 4885 (1991).
[34] T. Kuhn and F. Rossi, Phys. Rev. B **46**, 7496 (1992).
[35] F. Rossi and T. Kuhn, Rev. Mod. Phys. **74**, 895 (2002).
[36] D. Ferry, A. Kriman, H. Hida, and S. Yamaguchi, Phys. Rev. Lett. **67**, 633 (1991).
[37] H. Hida, S. Yamaguchi, A. Kriman, and D. Ferry, Semicond. Sci. Technol. **7**, 8154 (1992).
[38] M. Nedjalkov, I. Dimov, and H. Haug, Phys. Status Solidi B **209**, 109 (1998).
[39] C. Ringhofer, M. Nedjalkov, H. Kosina, and S. Selberherr, SIAM J. Appl. Math. **64**, 1933 (2004).
[40] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. A. Quaranta, Phys. Rev. B **12**, 2265 (1975).
[41] H. Shichijo and K. Hess, Phys. Rev. B **23**, 4197 (1981).
[42] M. Fischetti and S. Laux, Phys. Rev. B **38**, 9721 (1988).
[43] D. Chattopadhyay and H. Queisser, Rev. Mod. Phys. **53**, 745 (1981).
[44] N. Sano, T. Aoki, M. Tomizawa, and A. Yoshii, Phys. Rev. B **41**, 12 122 (1990).
[45] M. Fischetti, Phys. Rev. B **44**, 5527 (1991).
[46] S. Bosi and C. Jacoboni, J. Phys. C **C9**, 315 (1976).
[47] P. Lugli and D. Ferry IEEE Trans. Electron Devices **ED-32**, 2431 (1985).
[48] N. Takenaka, M. Inoue, and Y. Inuishi, J. Phys. Soc. Jpn. **47**, 861 (1979).
[49] D. Ferry, R. Akis, and D. Vasileska, *Intl. Electron Devices Meeting* (IEEE Electron Devices Society), 287 (2000).
[50] L. Shifren, R. Akis, and D. Ferry, Phys. Lett. A **274**, 75 (2000).
[51] S. Ahmed, C. Ringhofer, and D. Vasileska, J. Comp. Electron. **2**, 113 (2003).
[52] C. Ringhofer, C. Gardner, and D. Vasileska, Int. J. High Speed Electron. Syst. **13**, 771 (2003).
[53] S. Haas, F. Rossi, and T. Kuhn, Phys. Rev. B **53**, 12855 (1996).
[54] M. Nedjalkov, I. Dimov, F. Rossi, and C. Jacoboni, Math. Com-

put. Modell. **23**, 159 (1996).
[55] H. Lee and M. Scully, J. Chem. Phys. **77**, 4604 (1982).
[56] K. L. Jensen and F. A. Buot, IEEE Trans. Electron Devices **38**, 2337 (1991).
[57] H. Tsuchiya and T. Miyoshi, *Proceedings of 6th International Workshop on Comp. Electronics* IWCE6, Osaka, p. 156 (1998).
[58] R. Sala, S. Brouard, and G. Muga, J. Chem. Phys. **99**, 2708 (1993).
[59] H. Tsuchiya and U. Ravaioli, J. Appl. Phys. **89**, 4023 (2001).
[60] C. Jacoboni, VLSI Des. **8**, 185 (1998).
[61] M. Pascoli, P. Bordone, R. Brunetti, and C. Jacoboni, Phys. Rev. B **58**, 3503 (1998).
[62] L. Shifren and D. Ferry, Physica B **314**, 72 (2002).
[63] L. Shifren and D. Ferry, J. Comp. Electronics **1**, 55 (2002).
[64] D. Ferry, S. Ramey, L. Shifren, and R. Akis, J. Comp. Electronics **1**, 59 (2002).
[65] L. Shifren and D. Ferry, Phys. Lett. A **285**, 217 (2001).
[66] L. Shifren, C. Ringhofer, and D. Ferry, IEEE Trans. Electron Devices **50**, 769 (2003).
[67] L. Shifren and D. Ferry, Phys. Lett. A **306**, 332 (2003).
[68] E. Merzbacher, *Quantum Mechanics*, Second Edition (Wiley, New York, 1970), p. 542.
[69] Simple examples are particle splitting/gathering and self-scattering widely used in the classical Monte Carlo method.
[70] P. Bordone, A. Bertoni, R. Brunetti, and C. Jacoboni, Math. Comput. Simul. **62**, 307 (2003).
[71] We consider the subspace of points $(\mathbf{k}',\mathbf{r}')$ having finite t_b. Fortunately this is the relevant sub-domain for the integral (16), since in the complementary subspace the integrand vanishes.
[72] C. Jacoboni, P. Poli, and L. Rota, Solid-State Electron. **31**, 523 (1988).
[73] C. Jacoboni, *Int. Electron Devices Meeting* (IEEE Electron Devices Society, Washington, DC, 1989), p. 469.
[74] L. Rota, C. Jacoboni, and P. Poli, Solid-State Electron. **32**, 1417 (1989).
[75] M. Nedjalkov and P. Vitanov, Solid-State Electron. **32**, 893 (1989).
[76] M. Nedjalkov and P. Vitanov, Solid-State Electron. **33**, 407 (1990).
[77] H. Kosina, M. Nedjalkov, and S. Selberherr, IEEE Trans. Electron Devices **47**, 1898 (2000).
[78] C. Jacoboni and L. Reggiani, Rev. Mod. Phys. **55**, 645 (1983).
[79] T. Schmidt and K. Moehring, Phys. Rev. A **48**, R3418 (1993).
[80] M. Nedjalkov, R. Kosik, H. Kosina, and S. Selberherr, *Proc. Simulation of Semiconductor Processes and Devices* (Business Center for Academic Societies, Japan, 2002) p. 187.
[81] C. Wordelman, T. Kwan, and C. Snell, IEEE Trans. Comput.-Aided Des. **17**, 1230 (1998).
[82] We note that there are always two states generated: $p^{\pm}(\mathbf{r}',0)=0$ since V_w is antisymmetric.
[83] H. Kosina, M. Nedjalkov, and S. Selberherr, *4th IMACS Seminar on Monte Carlo Methods, MCM-2003*, Monte Carlo Methods and Applications (Berlin, 2003), p. 6.
[84] The analogy with the Dyson's equation used in the theory of Green's functions is straightforward.
[85] H. Kosina, G. Klimeck, M. Nedjalkov, and S. Selberherr, *Proceedings Simulation of Semiconductor Processes and Devices* (Boston, 2003), p. 171.
[86] H. Tsuchiya, M. Ogawa, and T. Miyoshi, IEEE Trans. Electron Devices **38**, 1246 (1991).
[87] H. Kosina, M. Nedjalkov, and S. Selberherr, J. Comp. Electronics **2**, 147 (2003).
[88] This equality can be easily seen if one considers the gradient expansion of B around the boundary $B=0$.
[89] I. Sobol, *Numerical Methods Monte Carlo* (Nauka, Moskow, 1973).