Scaled Ensemble Monte Carlo

A. M. Kriman[1] and R. P. Joshi[2]

[1]Dept. of Elec. and Comp. Engineering,
State University of New York at Buffalo,
Buffalo, New York 14260 USA

[2]Dept. of Elec. and Comp. Engineering,
Old Dominion Univ.,
Norfolk, Virginia, 23529 USA

Abstract

We introduce a scaled ensemble Monte Carlo (SEMC) technique, useful for obtaining statistically significant results far into the low-density tails of carrier distributions. Standard approaches for studying this regime (weighted EMC) are based on determining single-particle distribution functions using a nonuniform trajectory sampling; the new technique instead stimulates a scaled distribution, with an energy-dependent scaling factor, but uses ordinary EMC weighting. This technique is flexible and simple to code. We display some results for bulk GaAs, with densities ranging over seven orders of magnitude, using only 10,000 particles.

I. General System Description

Our treatment is in the classical regime, in the absence of significant interparticle correlations, so the system is described formally by the usual time-dependent single-particle distribution function $f(\mathbf{r},\mathbf{p};t) = f(\mathbf{x};t)$ [$\mathbf{x} = (\mathbf{r},\mathbf{p})$ is a phase space coördinate]. The distribution function obeys the Liouville equation:

$$\frac{\partial f}{\partial t} = \{H,f\} + \left(\frac{\partial f}{\partial t}\right)_{\text{coll}}, \tag{1}$$

where $H = H(\mathbf{x})$ is the time-independent single-particle Hamiltonian, and $\{\bullet,\bullet\}$ is the Poisson bracket. We make the usual assumptions that collisions or scattering events take place on time scales much shorter than the time between collisions, so effects such as collision broadening can be neglected. Further, the scattering events are in fact approximated as instantaneous, so that intracollisional field effects can also be ignored. Under these assumptions, all important sources of potential and phonon scattering can be written in the form

$$\left(\frac{\partial f(\mathbf{x};t)}{\partial t}\right)_{\text{coll}} = -\Gamma(\mathbf{x};t)\, f(\mathbf{x};t) + \int \Gamma(\mathbf{x},\mathbf{x}';t) f(\mathbf{x}';t)\, d\mathbf{x}' \ . \tag{2}$$

II. Scaling Formalism

The approach developed depends primarily on the observation that an energy- and time-dependent scaling factor commutes with noncollisional term on the right-hand side:

$$s(H,t)\ \{H,f\} = \{H,s(H,t)f\} \ . \tag{3}$$

Thus, we define a *scaled* distribution function defined by

$$\hat{f}(\mathbf{x};t) \equiv s(H(\mathbf{x}),t)\ f(\mathbf{x};t) \ . \tag{4}$$

The principal advantage of multiplying by an energy-dependent factor arises from counting statistics. If we are interested in the distribution function in some region about the phase space point x, we consider as a function of time the number of simulation particles N in the vicinity of that point. The fractional error in $\hat{f}$ is then $1/\sqrt{N}$, which is the fractional error in $f \equiv s^{-1}\times\hat{f}$ as well. For regions of low phase-space density, N is proportional to $f(\mathbf{x};t)$. Thus, for example, to examine regions where the phase space density is down by seven orders of magnitude from the maximum, one needs many times of ten million simulation particles to keep the error from exceeding the estimate.

A standard solution to this problem — weighted EMC — is to define a region of interest in phase space, and to perform multiple simulations of those few particle trajectories which enter it. Usually, this is coded in a way that sharply distinguishes high- and low-density regions. If the density falls smoothly, there is no efficient place to draw the boundary between these two regions. Other, more subtle approaches have also been used [1].

The present approach is based upon scaling the distribution function that is simulated by EMC, rather than upon a weighting the EMC simulation of an unscaled distribution function. The choice of s determines the trajectory density. In principal, these two approaches may be equivalent in particular cawses. However, a weighted EMC elimiinates the usual identification between individual initial particles and individual trajectories sampled (in what is the Monte Carlo integration of the Boltzmann equation). As a result, one loses the intuitive simplicity of regarding sampled trajectories as individual particles of a large ensemble. In a scaled EMC, on the other hand, one preserves a one-to-one correspondence of initial condition to trajectory, and it remains possible to regard the trajectories sampled as the actual paths of individual particles. As we describe below, however, in order to redistribute the statistical sampling weight, one pays the price that the trajectories do not follow the paths of ordinary particles.

III. Time-dependent scaling

Particle-number conservation imposes an important constraint on how $s(H(\mathbf{x}),t)$ is allowed to be chosen. By appropriate normalization, the total number $\hat{N}$ of simulation particles in the simulation of $\hat{f}$ is the integral of the scaled distribution function:

$$\hat{N}(t) = \int \hat{f}(\mathbf{x})\, d\mathbf{x} = \int s(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x} \,. \tag{5}$$

If s is chosen to emphasize high-energy regions which have low density, then during a relaxation, thermalization will cause a transfer of (real) electrons to lower-energy regions where s is smaller. If s is not allowed a time-dependence to compensate, the total number of simulation particles of $\hat{f}$ must decrease — degrading the statistics in $\hat{f}$. Conversely, a heating mechanism would increase the number of simulation particles, improving statistical precision but possibly requiring undesirable computational expense. By allowing s to have a time-dependence, we accomplish in the time domain what the energy dependence of s accomplishes for the energy domain: reduce variations in particle number so that fractional errors can be kept at an acceptable level throughout the region of interest, with the least computational effort.

We chose a simple form for the joint energy-and-time dependence of s: we let them be independent factors. This can be written

$$s(H,t) = \exp\big(\alpha(t)+\gamma(H)\big) \,. \tag{6}$$

Furthermore, we let the function $\gamma(H) = -H / k_B T^{eff}$. This is appropriate for distributions f which are approximately characterized by effective temperatures below T^{eff} [Higher temperatures lead to normalization problems with $\hat{f}$.] This choice is also convenient

computationally: a linear function γ implies that inelastic scattering rates as well as elastic rates, are space-position-independent, and can be stored efficiently in look-up tables.

$\hat{f}$ then obeys a modified Liouville equation

$$\frac{\partial \hat{f}}{\partial t} = \{H,\hat{f}\} + \left(\frac{\partial \hat{f}}{\partial t}\right)_{\text{coll}}, \tag{7}$$

in which the "collision" or scattering term is defined by

$$\left(\frac{\partial \hat{f}}{\partial t}\right)_{\text{coll}} - \frac{1}{s}\left(\frac{\partial s}{\partial t}\right)\hat{f} = s\left(\frac{\partial f}{\partial t}\right)_{\text{coll}}. \tag{8}$$

Equations (2) and (8) specify completely the modified scattering term $(\partial \hat{f}/\partial t)_{\text{coll}}$. However, in order to implement a Monte Carlo time-evolution, one must determine scattering rates for the scaled problem which are analogous to the out-scattering rates $\Gamma(x;t)$ and the in-scattering rates $\Gamma(x,x';t)$ of the unscaled problem. There is some freedom in way this is done. One well-known degree of freedom is associated with self-scattering:

$$\left\{ \begin{array}{ll} \Gamma(\mathbf{x};t) & \rightarrow \Gamma(\mathbf{x};t) + \Delta\Gamma(\mathbf{x};t) \\ \Gamma(\mathbf{x},\mathbf{x}';t) & \rightarrow \Gamma(\mathbf{x},\mathbf{x}';t) + \Delta\Gamma(\mathbf{x};t)\,\delta(\mathbf{x}-\mathbf{x}') \end{array} \right\}, \tag{9}$$

where ordinarily $\Delta\Gamma$ is chosen to make the total out-scattering a positive constant. This is a kind of gauge transformation, in which the physically-significant total scattering rate is fixed, while unobservable components of the in- and out-scattering rates make the numerical implementation tractable. A more general kind of gauge transformation is made in the scaled EMC approach, leading to off-diagonal in-scattering rates of the form

$$\hat{\Gamma}(\mathbf{x},\mathbf{x}') = G(\mathbf{x},\mathbf{x}') + G(\mathbf{x})\frac{1}{\hat{N}}\hat{f}(\mathbf{x}) \quad \text{for } \mathbf{x} \neq \mathbf{x}'. \tag{10}$$

The second term on the right-hand side leads to a kind of attractive interparticle scattering. This performs a rôle similar to that of trajectory iteration in weighted-EMC approaches: simulation particles entering critical regions are given greater weight, and are effectively caused to perform multiple traversals. However, in SEMC this weighting is implemented smoothly, rather than abruptly at the boundary of a region of interest, and it is accomplished with a fixed number of particles undergoing essentially ordinary scattering.

IV. Simulations

We have applied the SEMC technique to bulk GaAs semiconductor at 300 K. We used parameters (deformation potentials, phonon energies, band structure, etc.) that have been confirmed empirically in previous simulations [2]. We have specifically neglected hot phonon effects and Coulomb scattering, so we have essentially modeled the electrons in intrinsic GaAs with weak laser excitation. The system was allowed to relax for 10 ps from the initial laser excitation, and the result plotted below were obtained as an average of the distribution funcrion during the last 0.5 ps (i.e., an average was performed of results of the last 100 5-ps observation times). Because of the averaging procedure used, there is a kind of local averaging of the density, so that the standard deviation of the results from a smooth fit underestimates the uncertainty of the simulation results.

Figure 1 shows the density as a function of energy of electrons in the central (Γ) valley. An ordinary (unscaled, unweighted) EMC simulation is compared with the scaled simulation using $T^{\text{eff}} = 400$ K. *Both of the simulations use the same (constant) number of simulation particles*: 10^4. In this valley, the SEMC simulation appears to be *less* smooth (i.e., to have greater

statistical uncertainty) at low energies. This is in fact correct: the price we pay for increasing the number of simulation points at higher energies is to reduce the number at low energies, wirh correspondingly poorer statistics. At higher energies, when the densities become smaller than 10^{-4}, the usual EMC becomes completely unreliable, since results correspond to single simulation particles in a bin. The results of the two EMC simulations are consistent (in the sense that their difference is not statistically significant), as they must be, since they are mathematically equivalent in the large-*N* limit. The equivalence at high energies is clearly due to the large error bars in the unscaled EMC (not shown) which completely cover the range of densities to zero. There is the *appearance* that the unscaled EMC predicts a systematically higher density at high energy, since all of the high-energy points plotted are above the corresponding SEMC points plotted. There are two reasons for this appearance. One is the local (time) averaging discussed, which has the effect of smoothing a random statistical fluctuation into an apparent systematic one. The other is that on the semilogarithmic scale, the points where EMC predicts zero density (and lie *below* the SEMC simulation) cannot be plotted.

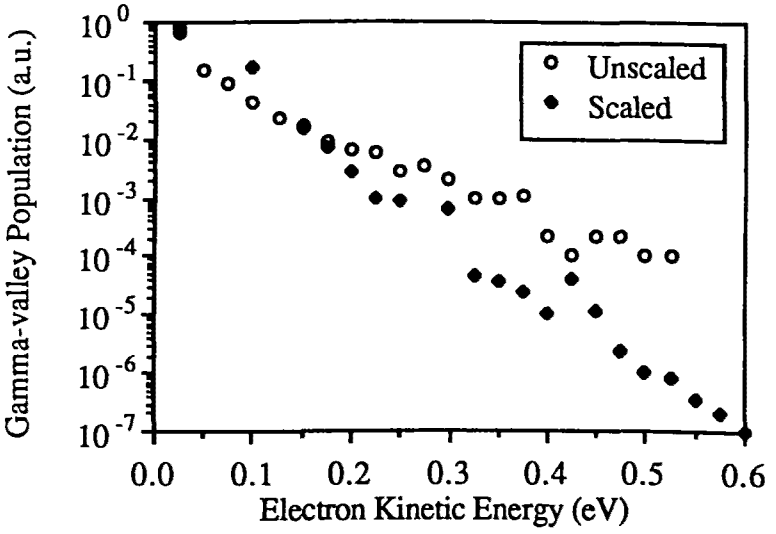

Figure 1. Central conduction valley population density.

Figures 2 and 3 show the same comparison as figure 1, for the L- and X-valleys, respectively. The thresholds at zero kinetic energies correspond to satellite band minima many times k_BT above the Γ-valley minimum, and so the regime of poor statistcs is reached more quickly. The SEMC in each case shows the smooth exponential fall-off of a nearly equilibrated system.

V. Conclusion

We have shown that a scaled EMC simulation can be implemented to significantly improve the statistics over broad ranges of low real densities.

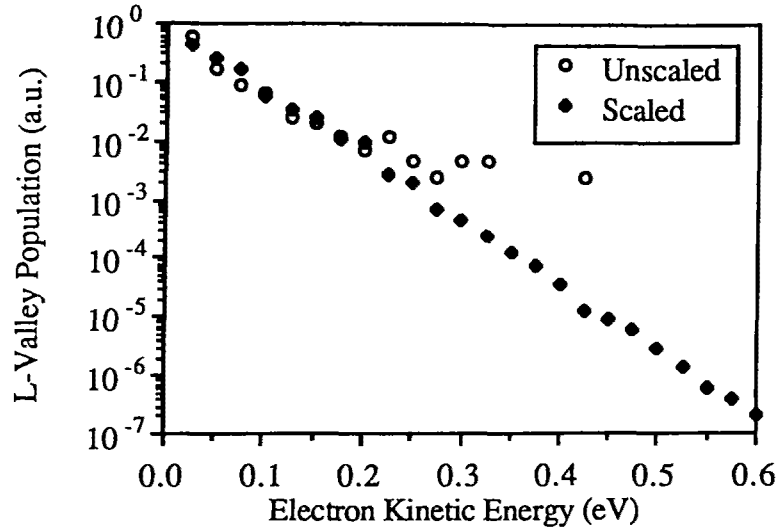

Figure 2. L-valley population density.

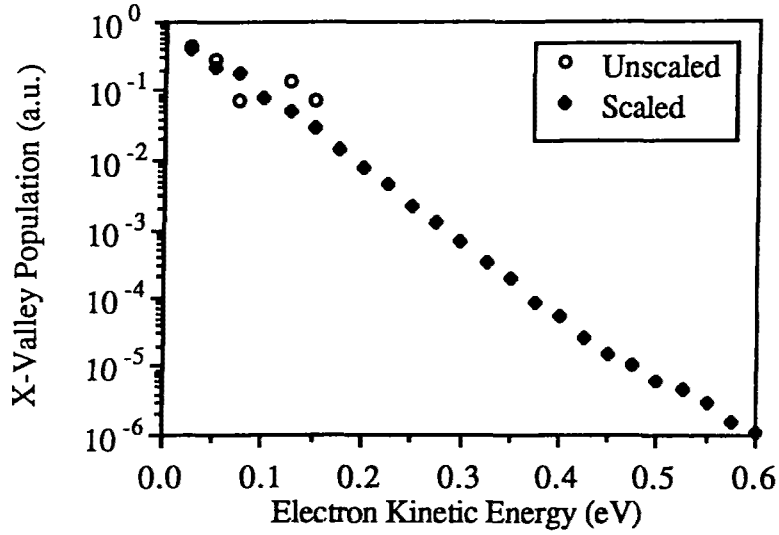

Figure 3. X-valley population density.

References

[1] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*, Springer-Verlag, 1989.

[2] M.-J. Kann, A. M. Kriman and D. K. Ferry, "Effect of Electron-Electron Scattering on Intervalley Transition Rates of Photoexcited Carriers in GaAs," Phys. Rev. B Vol. 41, No. 18, pp. 12659-12665, 15 June 1990.