A Self-Consistent Monte Carlo Simulator for Deep Submicron MOSFETs

F. Venturi, R. K. Smith, E. Sangiorgi, M. R. Pinto and B. Ricco'†

AT&T Bell Laboratories, Murray Hill, NJ 07974
† University of Bologna, Bologna, Italy

An efficient, self-consistent Monte Carlo (MC) simulator has been developed suitable for general silicon devices, including those with regions of high doping/carrier densities. Key features include an original iteration scheme between MC and Poisson's equation and an almost complete vectorization of the program. The simulator has been used to characterize nonequilibrium effects in deep submicron nMOSFETs. Substantial effects are noticeable at gate lengths of $0.25\mu m$ at room temperatures.

Nonequilibrium transport effects in very small geometry silicon MOSFETs have recently been reported; c.f. [1,2]. Most of this work has been based on measurements of transconductance (g_m) which do not seem to be affected unless extremely short channel lengths ($\approx 0.1\mu m$) and/or low operating temperatures are used. Without performing a more detailed analysis, it still remains unclear to what extent and at what dimension nonequilibrium effects are important. A Monte Carlo based (MC) device simulator is well suited to perform such a task since it treats the Boltzmann transport equation directly. It is to be noted that the carrier distributions obtained via MC *must* be computed self-consistently with the electrostatic potential as demonstrated by a comparison of surface densities for a MOSFET with a gate (mask) length $L_{gate} = 0.25\mu m$ (fig.1). An unfortunate result then is that the highly doped source and drain regions - required in order to provide equilibrium boundary conditions - would seem to force an inordinate amount of computation time [3], implying that MC may be impractical.

We report here on the development of a new, computationally efficient self-consistent MC simulator for n-MOSFETs. Given a solution from a conventional drift-diffusion (DD) simulator (e.g. [4]) as an initial guess, the program solves iteratively the carrier transport equation by means of the MC method and Poisson's equation, until convergence is achieved. Several improvements to the originally proposed scheme [5] have been made in order to simulate real devices in a reasonable amount of time. These are summarized below and will be described in some detail in the presentation:

- Since the usual self-scattering (SS) scheme is very inefficient due to the wide range of variation in the scattering probabilities, a multi-threshold SS scheme has been introduced, which reduces the number of SS events to about 10% of the total.
- To reduce the statistical noise in the electron distribution in poorly populated regions like the channel (compared to source or drain) or the high energy tails, a sample multiplication technique is used to magnify the rare particles while preserving the total charge.
- The direct solution of the linear Poisson's equation (holding the concentrations fixed) is very sensitive to noise in the densities in the highly populated regions. We instead solve Poisson's equation nonlinearly using the quasi-Fermi level provided by the MC simulator. In this way, the statistical fluctuations in the potential are reduced due to the logarithmic dependence of both the potential and the quasi-Fermi level on the densities.
- To accelerate convergence without losing stability, the electron temperature obtained from the MC data rather than the lattice temperature is used to scale the potentials in the Poisson solver. In the high electron temperature regions, where the MC and DD solutions differ most, this scheme allows the potential to change by a larger amount; on the other hand, in the source and drain regions, where most of the electrons are cold, the lower allowed variation keeps the solution stable. Fig.2 shows the improvement in the rate of convergence using this technique for a point at the end of the channel.
- By grouping particles according to their processing status, we are able to simultaneously treat multiple particles, achieving a rate of approximately 10^5 scattering events per second of Cray XMP (single processor) cpu time. As a result, the time required for the self-consistent simulation of a bias point is about 30 min.

The simulator has been used to analyze deep submicron MOS devices at T=300K with L_{gate} varying from 0.15-0.75μm. Consistent with recent experimental reports, nonequilibrium effects do not seem to be observable in the computed g_m down to $L_{gate} \approx 0.25\mu m$; at $0.15\mu m$, the MC g_m is more than 2 times the DD result. Figs. 3a-c show the average electron velocities in the channel for $0.75\mu m$, $0.25\mu m$ and $0.15\mu m$ devices respectively as computed using MC and DD. Interestingly both the $0.25\mu m$ and $0.15\mu m$ devices exhibit MC velocities substantially in excess of that predicted by DD throughout most of the channel. However, at the source end of the $0.25\mu m$ FET, the two velocities are nearly identical, leading to the same g_m. On the other hand, in order that current continuity be maintained, the larger average velocity throughout the rest of the channel implies lower inversion charge (fig. 1) and hence smaller gate capacitance and faster switching speeds. Fig. 4 illustrates this point by showing a plot of the transistor cutoff frequency f_T as a function of L_{gate} as predicted by MC and DD; both intrinsic and overlap capacitance is included. At $0.25\mu m$, the MC prediction is 20% higher than DD demonstrating conclusively the need for a self-consistent, nonequilibrium treatment of carrier transport for future MOSFET design.

[1] G. G. Shahidi, et. al, *IEDM Tech. Dig.*, p. 824, 1986.
[2] M.-C. Jeng, et al., *IEDM Tech. Dig.*, p. 710, 1987.
[3] T. Kobayashi, et al., *IEEE Trans. Elect. Dev.*, vol. ED-32, p. 788, 1985.
[4] M. R. Pinto, et al., *PISCES-IIB*, Stanford University, 1985.
[5] C. Jacoboni, et al., *Rev. of Mod. Phys.*, vol. 55, p. 645, 1983.
[6] C. Hao, et al., *Solid State Elect.*, vol. 28, p. 733, 1985.

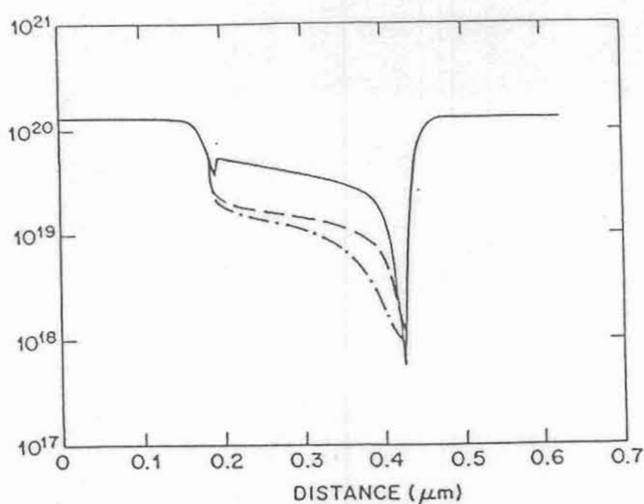

Fig.1: Surface densities in a 0.25μm device at $V_{gs} = 1.5\,V$ and $V_{ds} = 2.0\,V$ in the DD (—), non-self-consistent MC (- -), and self-consistent MC (— •) case.

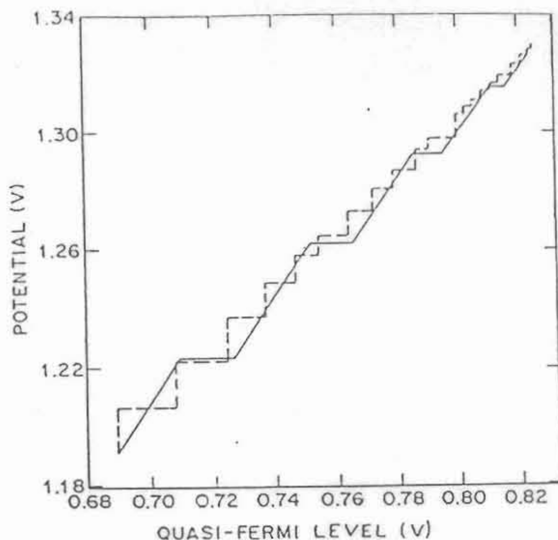

Fig.2: Achievement of the convergence for a point using the lattice (- -) or the electron (—) temperature in the Poisson solver. Each horizontal segment represents a MC iteration; each vertical (or diagonal) segment corresponds to a Poisson iteration.

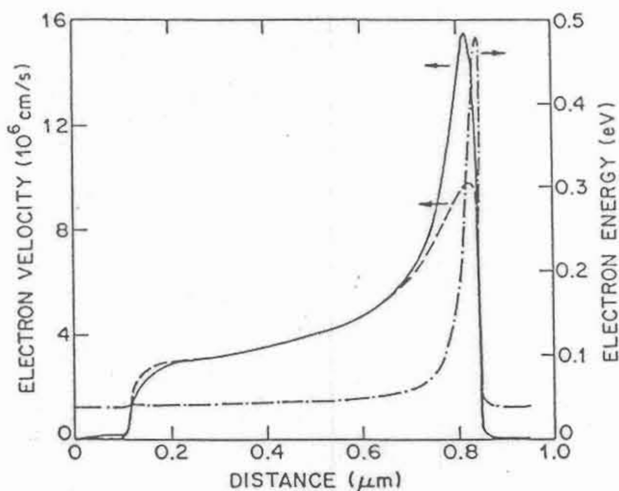

Fig. 3a

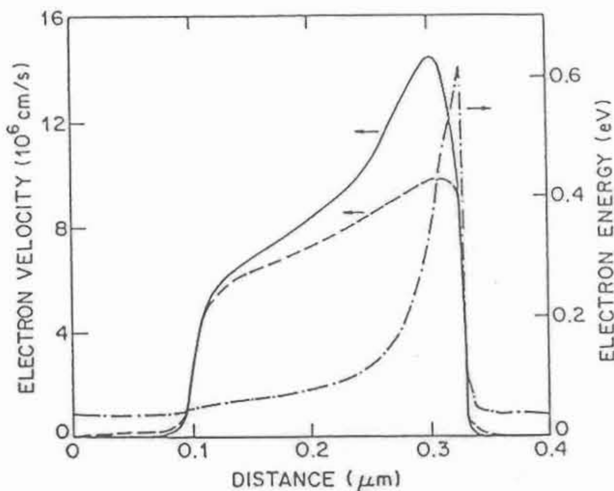

Fig. 3b

Fig.3: Average electron velocity in the MC (—) and DD (- -) simulations and surface electron energy in the MC (— •) solution for $V_{gs} = 1.5\,V$ and $V_{ds} = 2.0\,V$ for $L_{gate} = 0.75\mu m$ (a), $0.25\mu m$ (b), $0.15\mu m$ (c). The DD mobility model includes the vertical and lateral field dependencies. The MC scattering models include the mechanisms described in [5] and a surface scattering model [6].

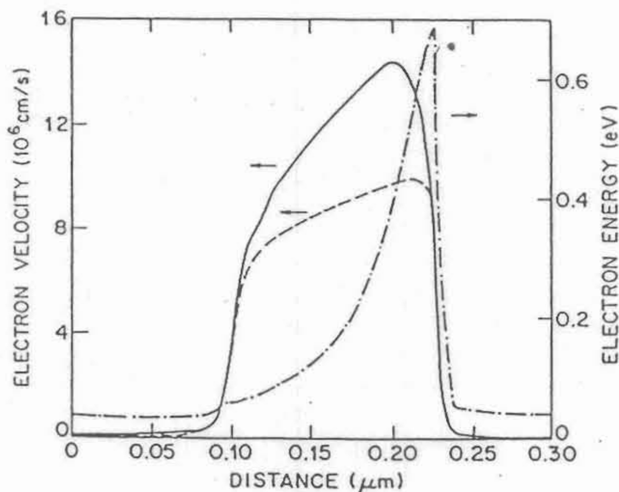

Fig. 3c

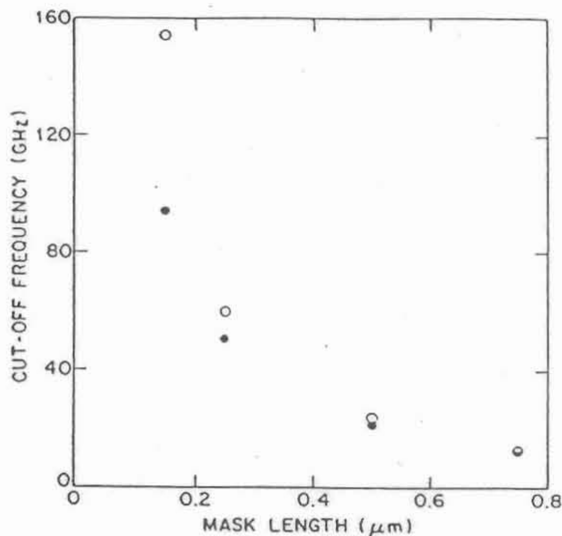

Fig.4: Cut-off frequency vs. gate mask length in the MC (o) and in the DD (•) simulation.