Thirty years of Monte Carlo simulations of electronic transport in semiconductors: Their relevance to science and to mainstream VLSI technology

Massimo V. Fischetti, Steven E. Laux, Paul M. Solomon, and Arvind Kumar
IBM SRDC, IBM Research Division, Thomas J. Watson Research Center,
P.O. Box 218, Yorktown Heights, New York, 10598
email: fischet@watson.ibm.com

At the 1996 ESSDERC one of us (MVF) gave a review of what were labeled "the first twenty years" of Monte Carlo simulations of electron transport in Si. At the time, emphasis was placed on their relevance to science – such as on the determination of the electron-phonon and impact ionization scattering rates, or on the suitable band-structure model to employ – but much less on their relevance to technology. Perhaps this was because such an impact on VLSI technology was, sadly, absent. Yes: Monte Carlo solutions of the Boltzmann Transport Equation had helped the improvement (and an improved understanding) of "moments" approximations, such as the drift-and-diffusion, energy-transport, and hydrodynamic models. But not much else... We should recall that, *at that time*, VLSI technology was evolving along the predictable evolutionary path of scaling Si MOSFETs. Electrostatic properties (*e.g.*, threshold-voltage and sub-threshold-swing) dictated by processing choices and accurately modeled by "mundane" Poisson solvers were the dominant effects. Concerns regarding electronic transport were, at best, marginal, hot-carrier effects being possibly the only exception.

"Scaling", as viewed then, is now gone, replaced by novel device designs (*e.g.*, FINFETs, Double-gate FETs) and novel active materials (*e.g.*, strained-Si, Ge, Si surfaces of "unusual" orientation) as technologist attempt to bypass the (real or perceived) end of this "conventional scaling". We do not have "pre-canned" drift-diffusion models of pair-production-*vs.*-field expressions for these materials, neither do we know the effect of novel gate-insulator materials on device performance. This quasi-revolution in VLSI technology had given Monte Carlo (MC) simulations a second wind and a practical importance. As it was the case a decade ago, science still benefits from the flexibility that the MC method allows, by decoupling physical models from numerical – mainly convergence – issues. Similarly, the reduced dimensions of the devices present interesting new physical aspects, such as (intentionally ignoring here "quantum concerns") the role of long-range Coulomb interactions and their effects on device performance (see Fig. 1). But the novelty of the structures and materials considered now in "real-world" VLSI technology requires the deep physical foundations of MC models (mainly: physically accurate band-structure, scattering matrix elements, and transport model valid all the way to the ballistic limit).

In the talk, having reviewed a few chapters of the science-related history of MC simulations, we shall delve into this "real world". Figure 2 represents a prototypical example of these accurate yet flexible physical foundations of MC models: Scaled MOSFETs using Si, Ge, InP and $In_{0.53}Ga_{0.47}As$ are simulated with a complete physical model. While at (relatively) large channel lengths the small transport mass of III-V compound semiconductors gives them some performance advantage, at the shortest channel length their small density-of-states effective mass makes them "choke" (too few "channels" -- *à la* Landauer – to carry current). Ge appears superior, but only on the chosen (111) surface orientations and ignoring band-to-band leakage issues. Additional examples – especially dealing with strained-Si devices, ballistic transport, and the meaning/relevance of the concept of "mobility" in small devices -- will also be given.

A full journal publication of this work will be published in the Journal of Computational Electronics.

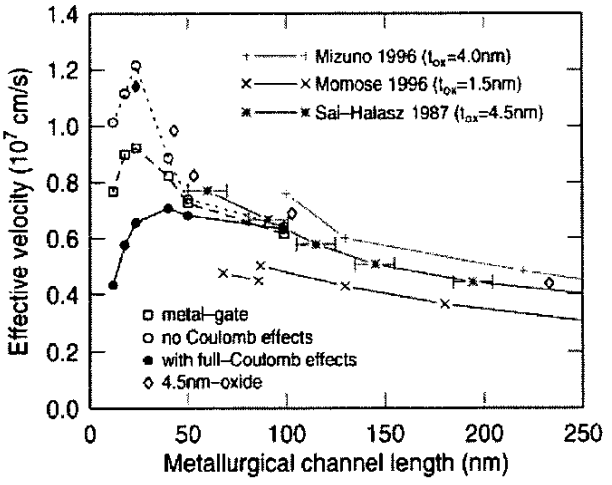

Figure 1. Effective velocity (transconductance divided by gate-to-source capacitance) for Si MOSFETs of various channel lengths with full Coulomb interactions (black solid symbols), with channel-gate interactions suppressed (as in a metal-gate device, black open squares) and without any Coulomb interaction. Clearly, interactions between electrons in the channel and those on the gate and source-drain regions depress the electron velocity.

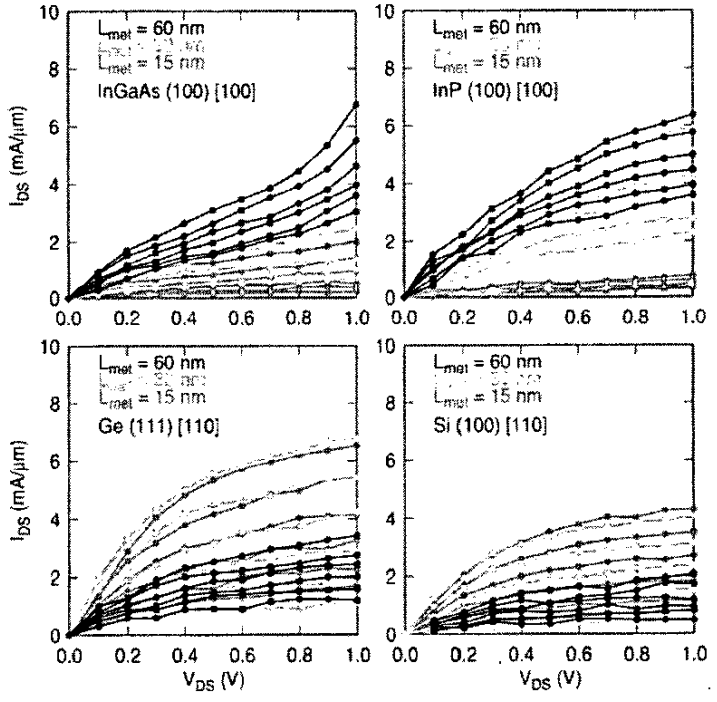

Figure 2. Drain-current vs. drain-to-source bias for MOSFETs of various (metallurgical) channel lengths built using four different semiconductors. Si-based devices exhibit a performance which improves when moving from channel lengths of 60 nm to 30 nm, saturating (or even losing performance, as in Fig. 1) at smaller dimensions. Ge devices built on the (111) surface and oriented along the [110] direction exhibit good performance, but degrade at the smallest channel length more than Si does (in addition to tunneling-leakage problems which we have not yet quantified). More dramatically, in InP- and InGaAs-based devices the small density-of-states mass is seen to represent a disastrous element at the smallest channel length.

A full journal publication of this work will be published in the Journal of Computational Electronics.