Cellular Automata for Device Simulation — Concepts and Applications

G. Zandler, M. Saraniti, A. Rein and P. Vogl

Walter Schottky Institute & Physics Department
TU – Munich, Germany

Abstract

We present a discussion of various concepts of cellular automata for semiconductor transport in the context of device simulation. A newly developed transformation for the kinetic terms of the Boltzmann equation into deterministic transition rules are found to be superior to probabilistic rules, allowing a complete suppression of statistical errors without any loss in numerical performance. To take advantage of the high speed of the resulting Cellular Automaton, a fast and flexible multigrid-solver for the Poisson equation has been developed. This enables us to study also fluctuations of transport quantities, which determine the high frequency noise behavior of MOSFETs, within the Cellular Automata approach. The reliability of the new CA approach for nanostructured devices is demonstrated by a study of gate length influence onto the drain current characteristics of a novel vertically grown MOSFET.

Cellular Automata for transport simulations in semiconductor devices: Concepts and recent developments

The physical effects involved in scaled down sub-μm devices require accurate physical simulation tools (1–3). Such simulation tools must be well beyond the level of the drift-diffusion approach to allow for a correct predictive description of high field transport. The Monte Carlo (MC) technique (4,5) is at present the most valuable approach to account for hot carrier effects and non-local transport phenomena typical for such devices. Unfortunately, it is also one of the numerically most costly methods and remained therefore limited to university and laboratory research.

Recently, a cellular automaton (CA) approach (6) has been developed as an efficient and discrete variant of the MC. In general, a cellular automaton consists of a lattice with a finite number of states attached to each lattice site. The realization of these states can be interpreted as a population with pseudoparticles. Their fictitious dynamics evolves on the microworld of the given lattice and can be updated simultaneously according to deterministic or nondeterministic rules in discrete time steps. Importantly, the dynamics of CA are governed by local rules, i.e. the updating of site variables involves only a small number of neighbors in each time step. For this reason, CA constitute one of the very few algorithms for physical processes which can optimally utilize massively parallel computer technology. In addition, the representation on a discrete lattice together with the locality of the dynamical rules allows an efficient and flexible treatment of complex geometries. The continuous physical quantities are obtained in practice by taking averages over many lattice sites.

In our present implementation for the solution of the Boltzmann equation, the microscale of the CA consists of a hexagonal two dimensional lattice in position space, each site of which has a finite number of momentum cells. These momentum cells are defined on the nodes of a periodic hexagonal close-packed lattice in three dimensional momentum space. The transition rules between these states, associated to collision events and semiclassical motion, are determined from the quantum mechanical scattering rates (in the same way as in MC) and from the classical equations of motion. Due to the locality in position space of quantum mechanical scattering events, which is a basic assumption underlying the Boltzmann equation, there is no principle problem to convert these transitions in momentum space into spatially local CA-rules (6). In a typical simulation, the number of momentum states are of the order of some 10⁴. The scattering rules are tabulated in coarse grained hierarchical tables to avoid excessive usage of memory. This results in moderate memory requirements for the pretabulated scattering rates, typically on the order of tens of megabytes for a standard subμM-MOSFET simulation. The hierarchical tables mentioned above allow the efficient determination of the transition to the final state.
While being of less importance for the quantum mechanical transition rules, a full three dimensional periodic momentum space is crucial for the determination of optimal transition rules originating in the kinetic terms of the Boltzmann equation. These drift and diffusion terms in the Boltzmann equation link the distribution function to its value in different position and momentum space locations, being therefore nonlocal in nature. Consequently, the crucial step in mapping the Boltzmann equation onto a cellular automaton, is the transformation of such terms into local transition rules for real and momentum space.

Two strategies, namely a transformation to probabilistic or alternatively, to deterministic transition rules, were developed in the past. In our first CA, we implemented a probabilistic treatment of the semiclassical dynamics of the pseudoparticles on the CA lattice. In such a scheme, a particle at lattice site $R$ stays in a given momentum state $K$ with a certain probability $P_K$ or changes its state to other momentum states with probability $1 - P_K$. Utilizing a periodic momentum state discretization, such hoppings can be restricted to nearest neighbor states by an appropriate choice of timestep $\Delta t$ and lattice constant $\Delta k$. The resulting scattering probabilities due to the electric field are given by $P_K = e \Delta t E(R)/(\hbar \Delta k)$, depending only on the momentum space lattice constant $\Delta k$, the field strength at lattice site $R$ and the chosen timestep ($\Delta t$). $e$ denotes the elementary charge and $\hbar$ the Planck constant. In close analogy, the real space dynamics is realized. Both dynamical rules for momentum and real space correspond to a random walk on a Markovian chain which is associated with a diffusion in momentum and real space respectively. As pointed out in Ref.(6), such artificial diffusion effects due to a probabilistic treatment of the particle trajectories are minimized for a periodic lattice. For a periodic momentum space lattice, $k$-space diffusion associated with this random walk can be calculated analytically for constant electric field in a lattice direction as $D_{\text{art}} = \Delta k^2 P_K (1 - P_K)/(2 \Delta t)$. Associated with this diffusion is an enhancement of the kinetic energy, the entropy and the diffusivity. In principle, this artificial diffusion can be reduced by a sufficiently small lattice constant $\Delta k$, however this becomes impractical in a three dimensional momentum space.

Alternatively, we proved that it is possible to transform the kinetic terms of the Boltzmann equation into a new deterministic hopping rule of the CA that completely suppresses this statistical error(8,9). The main point is to replace the probabilistic treatment of the kinetic terms by a deterministic discrete free flight. This procedure can be illustrated most easily in one dimensional momentum space. There, we introduce an additional internal degree of freedom, which is called color in the following. This color changes its value each timestep according to the local electric field strength. The k-state of a particle remains unchanged up to a critical color which corresponds to a change in momentum equal to the lattice constant $\Delta k$ in $k$-space and hops subsequently into its nearest neighbor state, resetting thereby its color. This procedure confines the statistical error to one $k$-cell, resulting in a vanishing artificial diffusion in momentum space. The restriction to nearest neighbors transforms the drift-term of the BE into a local interaction on momentum cells, in complete analogy to the treatment of the real-space diffusion-term of the BE (6). Importantly, we found that the new implementation of the CA does not require more computer time per iteration than our earlier two-dimensional implementation (6) even though it is significantly more accurate.

We found perfect agreement for all relevant stationary and transient ensemble averaged transport quantities with MC calculations, namely for mean velocity, average kinetic energy and diffusivity. Due to the complete suppression of artificial diffusion effects, CA simulations describe also correctly the fluctuations of transport quantities. We calculated identical autocorrelation functions for velocity and energy fluctuations in bulk silicon for different field strength from CA and MC simulations.

The numerical efficiency of our present CA implementation was tested on several computer architectures and compared with MC codes of similar complexity. With both programs we have simulated $10^4$ particles in bulk silicon on one vector-processor of a Cray YMP8/8-128. The resulting cpu time as a function of field strength is shown in Fig.1). Starting with a speed-up of 25 at low electric field strengths, the lattice gas method performs up to 40 times faster than the MC method at high fields. This increase in speed-up with increasing electric field can be attributed to the increasing number of scattering events, whose realization is less time consuming in the CA due to the possibility of precalculated scattering tables in momentum space. The speed-up at low fields, on the other hand shows the numerical efficiency of the deterministic transition rules for the free flight. The high numerical efficiency of the CA method on single processor machines is further supported by its ideal parallel
performance. On the parallel architectures of a Cray YMP (10) as well as of a PowerPc "Parsytec X"plorer" (11) we found a nearly linear scaling of computational time with the number of processors used. As discussed in Ref. (12) the reduced data exchange between processors required for the CA compared to MC allows to maintain the CA/MC speed up for the particle dynamics in device simulations, utilizing a geometrical decomposition of device geometries.

Applications to device simulation

Due to the high intrinsic speed of the CA algorithm, an efficient application of this method to self consistent device simulation requires a comparably fast Poisson solver. By applying standard Successive Overrelaxation (SOR) (13) techniques, the solution of the Poisson equation turns out to be the bottleneck of the entire simulation already in a single processor computer environment. For a typical two dimensional device geometry, an SOR-based solver takes about 90 percent of the overall simulation time. This bottleneck was resolved by the development of a fast and flexible Poisson solver (14), based on an iterative multigrid (MG) method. The speed improvement obtained by using the MG Poisson solver is more than one order of magnitude with respect to the SOR, even with unfavorable boundary conditions (13). This speed-up makes the solution of the Poisson equation in the simulation as fast as a CA step, thus removing the bottleneck due to the unbalanced ratio between CA and SOR. Applying the resulting CA device simulator to a highly doped submicron MOSFET, using $5 \times 10^5$ particles on a grid with

100x200 blocks for the solution of the Poisson equation, we end up with typical cpu times clearly below one hour.
on a HP 735/9000 workstation for one bias point on a I-V characteristics, which is already close to typical cpu times for approximative schemes based on hydrodynamic models. This numerical efficiency also allows us to simulate the long time sequences necessary to calculate correlation functions of fluctuations of terminal quantities, offering the possibility to perform microscopic noise studies (16) for realistic device geometries (17). The reliability of the CA approach was verified for various nanostructured devices (8,9,18,12,19). As a very recent application we present results for novel vertically grown silicon MOSFETs (15) with ultra short channels. The geometry of these devices is shown in fig.3a. To allow a comparison with the experiments taken from literature, we varied only the channel length without scaling doping concentration or oxide thickness. Such a comparison of the drain current characteristics for three gate lengths down to 65 nm is depicted in fig.3b. Pronounced short channel effects cause a strong increase of the drain output conductance with decreasing channel length. Besides electrostatic effects, velocity overshoot enhances the drain current as well as the output conductance by reducing the gate length. For a channel length of 65 nm, this results in an average channel velocity of more then twice the saturation velocity of silicon. All these effects are exactly accounted for in the CA method, demonstrating its capability to describe high field transport in modern submicron devices in a numerically efficient way.

Acknowledgment: This work was supported by the Siemens project SFE II and SFB 348

References