

3-D TCAD Monte Carlo Device Simulator: State-of-the-art FinFET Simulation

Gabriela F. Furtado¹, Vinícius V. A. Camargo², Dragica Vasileska³, Gilson I. Wirth¹

¹Programa de Pós-Graduação em Microeletrônica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

²Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, RS, Brazil

³School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, USA

e-mail: gabriela.furtado@ufrgs.br

Abstract—This work presents a comprehensive description of an in-house 3D Monte Carlo device simulator for physical modeling of FinFETs. The simulator was developed to consider variability effects properly and to be able to study deeply scaled devices operating in the ballistic and quasi-ballistic regimes. The impact of random dopants and trapped charges in the dielectric is considered by treating electron-electron and electron-ion interactions in real-space. Metal gate granularity is included through the gate work function variation. The capability to evaluate these effects in nanometer 3D devices makes the presented simulator unique, thus advancing the state-of-the-art. The phonon scattering mechanisms, used to model the transport of electrons in pure silicon material system, were validated by comparing simulated drift velocities with available experimental data. The proper behavior of the device simulator is displayed in a series of studies of the electric potential in the device, the electron density, the carrier's energy and velocity, and the Id-Vg and Id-Vd curves.

Index Terms—TCAD Simulation, Monte Carlo Method, FinFETs.

I. INTRODUCTION

The constant reduction of semiconductor device dimensions gave rise to an increase in variability, leading to randomly distributed device parameters and circuit response [1-3]. In other words, device scaling has made the impact of charge trapping at localized states in the gate dielectric or at the Si/SiO₂ interface grow significantly, thus turning these effects into relevant reliability concerns for modern semiconductor technologies [4, 5]. It has become, therefore, mandatory to develop proper models and simulators that are able to accurately predict the devices electrical behavior and their fluctuations due to variability issues. Device simulation may enable the prediction of device behavior and dependence on different parameters without the extensive need of experimental data for calibration, thus leading to reduction of manufacturing costs and to performance improvement in future technologies. In this regard, technology for computer-aided design (TCAD) tools can be very useful, allowing the user to access internal information of device parameters, and enabling the evaluation of different scenarios rapidly, thus decreasing design cycle time, and giving insight for the next trials on manufacture and shortening the time to market.

To address the market and academy needs, constant efforts have been made to implement fast and accurate device simulators. Generally, there are two main components of semiconductor device simulation that must be solved self-consistently with one another: transport equations and the electromagnetic fields.

The electromagnetic fields are driving forces for charge transport and arise from external sources, charge distribution or current densities. In general, the electric and magnetic fields can be obtained from the solution of Maxwell's equations, but, in certain conditions, it is only necessary to evaluate the quasi-static electric field resulting from the solution of the Poisson's equation.

The transport equations, alternatively, can be based on semi-classical or quantum transport models, the latter being more precise, but frequently prohibitively time-consuming. A variety of semi-classical transport models exist and utilize direct or approximate solution of the Boltzmann Transport Equation (BTE). The Monte Carlo method, that in the long time limit solves the Boltzmann Transport equation, has been used extensively for nonequilibrium transport calculation in semiconductor materials and devices in the past decades. It is a powerful tool to evaluate transport properties of carriers in both bulk semiconductors and for particle-based device simulations.

In this work, a 3D TCAD Monte Carlo state-of-the-art device simulator is presented. In the next sections, Monte Carlo technique will be discussed in detail and its application to device simulation will be presented. Section 2 introduces the Monte Carlo method for particle transport. The main scattering mechanisms for silicon, that are incorporated in our theoretical model and their corresponding scattering rates, are discussed in details in Section 3. The details of the particle-based device simulator are introduced in Section 4, where we show how particle transport is coupled to Poisson's equation. Simulation results from bulk silicon Monte Carlo algorithm regarding carrier velocity and energy over time for different electric fields are shown in Section 5. In Section 6 we present simulation results obtained with the 3D TCAD n-FinFET simulator, such as drain current dependence on gate and drain voltages, electric potential distribution, electron density and carrier velocity and energy over the channel length. Conclusive comments regarding this work are presented in Section 7.

II. MONTE CARLO METHOD FOR CHARGE TRANSPORT

The Monte Carlo (MC) technique is a stochastic method to solve large and complex mathematical problems; in this work, the MC method is applied to simulate nonequilibrium particle dynamics in semiconductor materials and devices, by solving the 3D BTE. The Monte Carlo algorithm consists of simulating free particle motions (free flights) terminated by random scattering events that repeatedly take place during the total observation time. The algorithm consists of

randomly choosing the duration of the free flight and then selecting stochastically the scattering event that aborts the free-flight motion and affects the final energy and momentum of the carrier. This procedure is constantly repeated, creating a free-flight/scattering loop that is interrupted only when the total simulation time is achieved.

A. Single-Particle Bulk Monte Carlo Method

The single-particle Monte Carlo method is applicable to transport analysis of bulk materials and generally tracks the motion of the carriers in the momentum space. The simulation evaluates the drift of a particle (free flight) under constant electric field during a randomly generated time called flight time. The free flight is terminated by a scattering event that is due to, for example, impurities or phonons. The sequence of free flight/scattering is repeated until the simulation end time is reached. The duration of the free flight depends on the total scattering rate, which is the sum of the scattering rates of each individual source of scattering. Fig. 1 shows a flowchart for the bulk Monte Carlo solver.

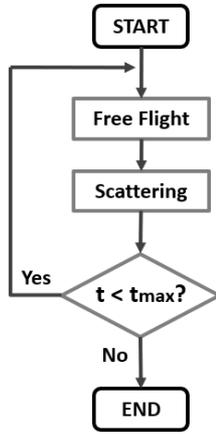


Fig. 1. Simplified Bulk Monte Carlo Flowchart.

In the initialization stage, the particle positions in k -space must be defined. The carrier energy is assumed to follow Boltzmann distribution, according to which

$$E = -\frac{3}{2} K_B T \ln(r_1), \quad (1)$$

where K_B is Boltzmann constant, T is the temperature and r_1 is a random number uniformly distributed between 0 and 1.

The magnitude of the k vector is determined by the dispersion (E - k) relation. Near the band minima, the bands are usually assumed to be parabolic, i.e., $E = \frac{\hbar^2 k^2}{2m^*}$. This approximation is accurate only for energies close to the valley minima. For higher energies, the parabolic band approximation becomes inaccurate, and a more complex, nonparabolic model must be employed. From the k , p perturbation theory, one arrives at the non-parabolic band approximation [6], according to which

$$k = \frac{\sqrt{2m^*E(1+\alpha E)}}{\hbar}. \quad (2)$$

In Eq. (2) the nonparabolicity factor α is a material dependent constant, given by $\alpha = \frac{1}{E_g} \left(1 - \frac{m_c}{m_0}\right)^2$, where m_c is the conductivity mass and m_0 is electron rest mass.

After the initialization of the particle energy and momentum, the next step in the simulation procedure is the generation of the free-flight time of each particle. The probability per unit time $P(t)$ of an electron moving freely during a time t and then scattering in the time interval dt is given by

$$P(t) = W_T[k(t)] \exp\left\{-\int_0^t W_T[k(t)] dt\right\} \quad (3)$$

where the total scattering rate $W_T[k(t)]$ is the sum of all energy-dependent scattering rates, being therefore itself dependent upon the electron energy. The free flight time t_r that precedes the scattering events can be determined by evaluating t for a certain $P(t)/W_T[k(t)]$ in Eq. (3). The integral in Eq. (3) cannot be solved analytically because of the complexity of $W_T[k(t)]$. With the introduction of a fictitious scattering event that leads to constant total scattering rate, this difficulty is overcome. This virtual scattering process is called self-scattering and it must have no impact on the k vector, being defined so that the electron state before and after the scattering event is the same. The self-scattering rate $W_{self}[k(t)]$ is chosen so that the total scattering rate $\Gamma = W_T[k(t)] + W_{self}[k(t)]$ becomes constant. Eq. (3) can now be rewritten as

$$P(t) = \Gamma e^{-\Gamma t}. \quad (4)$$

Assuming that $r = \frac{P(t)}{\Gamma}$ is uniformly distributed random number between zero and one, the free flight time t_r can be expressed as

$$t_r = -\frac{1}{\Gamma} \ln r. \quad (5)$$

Having the free-flight time defined, the carrier will drift during t_r . Since in the bulk Monte Carlo method the electric field is assumed to be constant during the carrier free-flight, the wave vector change Δk is calculated using Newton's law, i.e. $\Delta k = \frac{q}{\hbar} F t_r$, where q is the carrier charge, \hbar is the reduced Planck constant and F is the electric field.

Thus, during the drift process, the carrier accelerates due to the electric field, and this process stops when it undergoes a random scattering event. There are several random scattering events that may occur in semiconductors; for silicon, the main processes are acoustic phonon scattering, non-polar optical phonon scattering and Coulomb scattering.

The selection of the scattering mechanism that terminates the carrier free flight is chosen by a generation of a uniformly distributed random number r_2 which lies between 0 and Γ . For N scattering mechanisms incorporated in our theoretical model [1,2,...,N], the cumulative scattering rate of the n -th mechanism includes the scattering rate of all previous mechanisms, i.e., $W_n[k(t)] = \sum_{i=1}^n W_i[k(t)]$. The randomly generated number r_2 will then be confronted with $W_n[k(t)]$; the n -th scattering mechanism will be chosen when $W_{n-1}[k(t)] < r_2 \leq W_n[k(t)]$. Since the calculation of $W_n[k(t)]$ is time-consuming, it is convenient to tabulate the scattering rate of every scattering mechanism for different quantized energy levels that will be used in the simulation and organize them in a *scattering table*.

The particle energy and wave vector may change after the scattering event. If an acoustic phonon scattering has happened, the energy remains the same as before the event, since it is considered to be an elastic process. If a non-polar optical

phonon scattering occurs, the energy of the particle will change by the phonon energy.

Note that, if a full-band model is used, one needs to solve the time-independent Schrödinger wave equation for a periodic potential $V(\mathbf{r})$, of the form

$$\left[-\frac{\hbar^2}{2m_0} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (6)$$

in which \hbar is the reduced Planck's constant, ∇^2 is the Laplacian operator, $\psi(\mathbf{r})$ is the eigenfunction to be determined and E is the corresponding energy eigenvalue.

The transition rate from an initial state \mathbf{k} to a final state \mathbf{k}' can be evaluated by using Fermi's Golden Rule, which is based on the solution of the time-dependent Schrödinger equation and on the first-order, time-dependent perturbation theory [7]. For the j -th scattering mechanism, Fermi's Golden Rule is expressed as

$$\Gamma_j(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} |M_j(\mathbf{k}, \mathbf{k}')|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}} \mp \hbar\omega), \quad (7)$$

where $\Gamma_j(\mathbf{k}, \mathbf{k}')$ and $M_j(\mathbf{k}, \mathbf{k}')$ are respectively the transition probability between initial state \mathbf{k} and final state \mathbf{k}' and the matrix element for the j th scattering mechanism, $E_{\mathbf{k}}$ and $E_{\mathbf{k}'}$ are the initial and final state energies of the carrier and $\hbar\omega$ is the energy absorbed (upper sign) or emitted (lower sign) during the transition.

The matrix element of the perturbation potential between states \mathbf{k}' and \mathbf{k} is given by $M(\mathbf{k}, \mathbf{k}') = \langle \psi_{\mathbf{k}'} | H' | \psi_{\mathbf{k}} \rangle$, and each scattering event leads to a different matrix element. It can also be written as $\langle m, \mathbf{k}' | V_j(\mathbf{r}) | n, \mathbf{k} \rangle$ for transitions from state \mathbf{k} in band n to state \mathbf{k}' in band m , $V_j(\mathbf{r})$ being the scattering potential of the j -th process.

B. Bulk Ensemble Monte Carlo Method

The ensemble Monte Carlo method consists of successively applying the single particle bulk MC method for different carriers. Instead of following the motion of only one carrier for successive time steps, the ensemble Monte Carlo method follows a statistical ensemble of non-interacting carriers during each time step. Labeling the carriers present in the material from 1 to N , the method consists of observing, for example, the first carrier free-flight-scatter during a time interval Δt (that is called observation time). This procedure is repeated until all N particles have independently drifted during the time interval Δt . At every Δt , the average momentum and energy of the carriers in the ensemble is calculated.

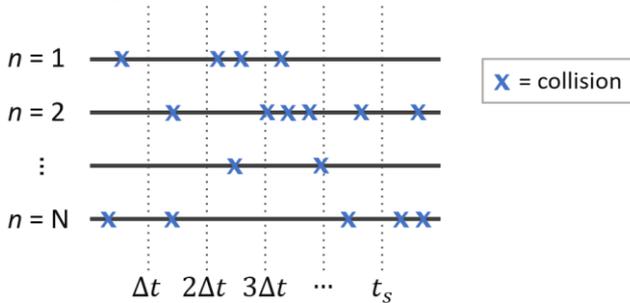


Fig. 2. Description of an Ensemble MC method. There are N carriers in the ensemble, Δt is the observation time and the total simulation time is t_s . Each blue symbol corresponds to a scattering event. [8].

Fig. 2 shows the drift of N particles during a time interval t_s . The observation time window is Δt , and the blue crosses indicate the moments when a scattering event takes place. Note that it is possible to have none or multiple scatterings events during one observation time interval.

III. SCATTERING MECHANISMS

Relevant scattering mechanisms for a silicon material system are: scattering due to intravalley acoustic and intervalley non-polar optical phonons, surface roughness scattering and Coulomb scattering. The electron-phonon scattering mechanisms are treated in momentum space. In particle-based device simulations, discussed later in the text, the Coulomb interactions, involving electron-ion and electron-electron interactions, are treated in real space using molecular dynamics approach. In silicon material system, two types of electron-phonon interactions are important: acoustic and non-polar optical phonon scattering. For the sake of completeness, their scattering rate equations will be discussed next, following the approach presented in [8].

The *elastic acoustic phonon scattering* rate out of state \mathbf{k} with energy E , for general nonparabolic bands (Fig. 3) is given by:

$$W(E) = \left(\frac{2\pi \Xi_d^2 K_B T_L}{\hbar \rho v_s^2} \right) \left(\frac{(2m_d^*)^3 \sqrt{E(1+\alpha E)}}{4\pi^2 \hbar^3} \right) (1 + 2\alpha E) \quad (8)$$

where Ξ_d is the deformation potential, T_L is the lattice temperature, ρ is the solid density, v_s is the velocity of sound in the crystal, $m_d^* = \sqrt[3]{m_l m_t^2}$ is the DOS effective mass and m_l and m_t are the electron longitudinal and transverse masses.

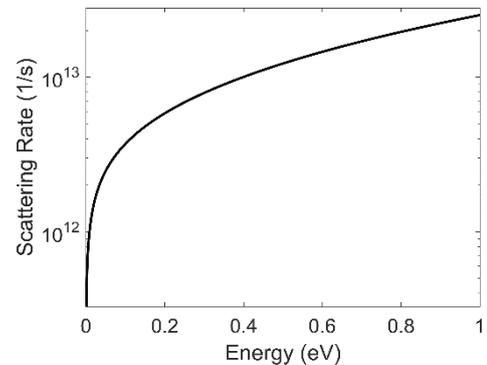


Fig. 3. Acoustic phonons scattering rate for silicon as a function of energy.

The *non-polar optical phonon scattering* rate (Fig. 4) is discussed next. Regarding the number of final states available for scattering (Z_j), the intervalley non-polar acoustic phonon scattering is divided in two categories, known as f-process and g-process. This division is based on the material's constant energy surfaces. If the scattering forces the carrier to move to a valley that is located along the original axis where the carrier was located previously, the carrier is scattering in a longitudinal direction, thus characterizing a g-process. In this case, there is only one valley available to scatter into, what makes $Z_j = 1$. Otherwise, if the carrier is scattered to a valley that is located in an axis different from the one where the initial valley is located, the carrier is scattered in a

transverse direction. There are, then, four options as feasible final valleys after the scattering, making $Z_j = 4$ for a f-process.

For nonparabolic bands, the zero-order approximation is given by

$$W(E) = \frac{\pi D_{ij}^2 Z_j}{\rho \omega_{ij}} \left[n(\omega_{ij}) + \frac{1}{2} \mp \frac{1}{2} \right] x \quad (9)$$

$$x \left(\frac{(2m_a^*)^3 \sqrt{E_f(1 + \alpha E_f)}}{4\pi^2 \hbar^3} \right) (1 + 2\alpha E_f)$$

where D_{ij} is the intervalley deformation potential, $\hbar\omega_{ij}$ is the energy of the non-polar optical phonon, E_k is the electron energy and ΔE_{ji} is the difference between the potential energy of valley j and the bottom of valley i , $E_f = E_k + \hbar\omega_{ij} - \Delta E_{ji}$ for phonon absorption and $E_f = E_k - \hbar\omega_{ij} - \Delta E_{ji}$ for phonon emission. The optical phonon energy $\hbar\omega_{ij}$ corresponds to the amount of energy the electron absorbs or emits when interacting with a phonon. For silicon, the energy difference ΔE_{ji} is zero, since the six valleys are equivalent.

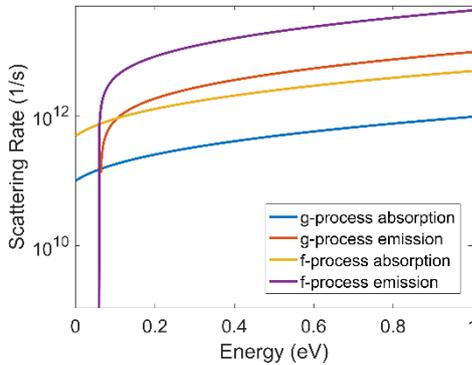


Fig. 4. Non-polar optical scattering rates for silicon as a function of energy.

Since the optical phonon energy $\hbar\omega_0$ is comparable to the carriers average thermal energy at room temperature, optical phonon scattering is considered inelastic. On the contrary, the energy of the acoustic phonons near the gamma point is very small (on the order of 1-2 meV) when compared to the thermal energy $K_B T$. This observation justifies the elastic approximation for acoustic phonons.

IV. DEVICE SIMULATION

In Section II, carrier transport in bulk semiconductors was evaluated by the numerical solution of the Boltzmann transport equations using the Monte Carlo method. However, in device simulation carrier transport and the driving fields are coupled and must be solved self-consistently. In order to simulate a device, therefore, the electric field that drives the charge transport must be taken into account by the solution of Poisson's equation. The electric field coming from the solution of Poisson's equation is responsible for the acceleration of the carriers that drift under the Monte Carlo phase, while the charge distribution itself defines the electric fields in Poisson's equation.

Over a very small time intervals, Poisson's equation may be decoupled from the BTE, thus simplifying the calculations. During this time interval, the carriers drift driven by

the frozen electric field that resulted from the solution of Poisson's equation at the end of the previous time interval. The Monte Carlo calculations lead to the charge distribution that, at the end of this time interval, will define the charge term in the Poisson's equation, which in turn provides the fields for the next Δt .

In addition to the potential calculation being obtained by the solution of Poisson's equation, instead of being spread in a infinite domain, the particles are restricted to move in a specific area defined by the device boundaries. These boundaries must be suitably defined into the code so that the particles may be reflected at the surface or may exit or enter the device through the terminals.

The structure of the Monte Carlo device simulation is presented next. Device simulation can be roughly divided into the following main blocks:

- Initialization of material parameters, device structure and carriers energy, momentum and position;
- Particle-mesh coupling;
- Solution of Poisson's equation;
- Interpolation of force to particle location.
- Monte Carlo free-flight/scattering;

Fig. 5 shows the device simulation framework based on the ensemble Monte Carlo method.

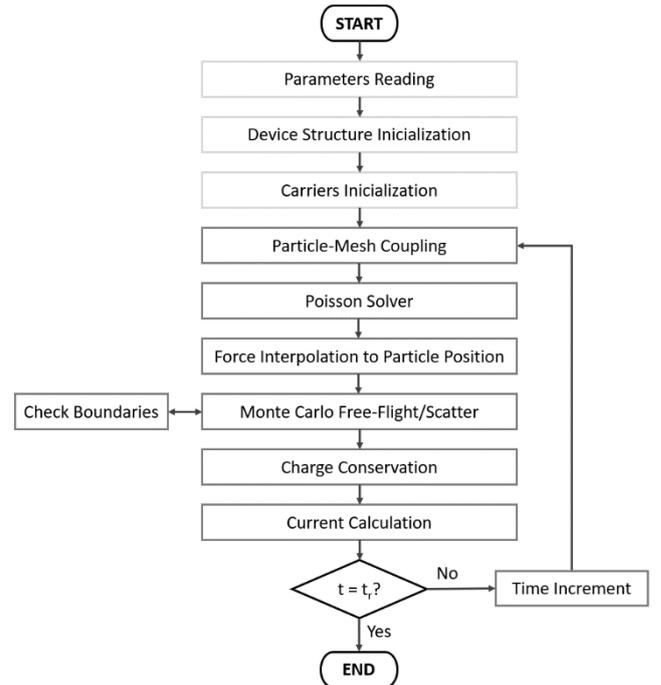


Fig. 5. Flowchart of a device simulator based on the ensemble Monte Carlo method.

As shown in Fig. 5, Monte Carlo device simulation can be divided in three main parts: calculating particle dynamics with appropriate boundaries conditions; verifying particles entering or exiting through the device terminals; and potential calculation with boundaries conditions [7].

A. Material Parameters Initialization

The required material parameters must be supplied to the simulator with an input file. Material parameters, such as

crystal density, velocity of sound in the crystal, dielectric constant and effective masses, are then read by the code.

B. Mesh size and Time-step

To allow for stable Monte Carlo device simulation, the mesh size (Δx , Δy and/or Δz) and the simulation time-step (Δt) must be chosen appropriately, correlating to each other in connection with the numerical stability. While Poisson's equation is solved at every Δt , the particle dynamics governed by the Monte Carlo algorithm takes place during this small time-step.

The mesh size defines the spatial resolution of the potential and is, therefore, determined by the charge variations. In order to be sensitive to these fluctuations, the mesh size must be smaller than the smallest wavelength of the charge variations, which is approximately equal to the Debye length (L_D). The mesh size in the simulation must be chosen to be smaller than L_D , so that relevant information on the potential profile is not lost. The mesh size can be changed by the user according to the device size - smaller devices can have a smaller grid, and larger devices can have proportionally bigger mesh size. The largest computational cost of a Monte Carlo 3D device simulator is due to the solution of the 3D Poisson's equation. The mesh size influences Poisson's equation solver directly, so it must be chosen wisely.

To be able to decouple Poisson's equation from the BTE, the simulation time-step must be suitably small. From the viewpoint of stability, this can be assured by making the time-step Δt smaller than the inverse of the plasma frequency.

Even though the mesh-size and the time-step were defined separately, they are indeed coupled; this can be checked by evaluating the maximum distance $l_{max} = v_{max} \Delta t$ that the carriers can move during Δt , where v_{max} is the maximum carrier velocity, approximated as the maximum group velocity of the carriers in the semiconductor. The time-step must be chosen to be small enough to make l_{max} smaller than the spatial mesh size estimated by taking Debye length into account. Since Poisson's equation is only solved at every Δt , the fields are updated only at these moments, and a too large time-step might lead to an unaccounted substantial change in the charge distribution and unphysical plasma oscillations.

C. Device Structure Description

The device must be described in the simulator by defining the boundary conditions, which enter the simulation in the Monte Carlo particle dynamics and in the numerical solution of the Poisson's equation. All design options are predefined, such as the device type (FinFET in this work) and the optional addition of characteristics like high-k dielectric, that affects the semiconductor device operation.

Except for the four FinFETs terminals and the oxide-semiconductor interfaces, the other surfaces are defined by reflecting boundary conditions. These boundaries are assumed to reflect perfectly the particles that reach them, reversing the velocity component normal to the surface when the carriers collide against the surface. The Neumann boundary condition of zero electric field in the direction normal to the surface is then applied to the Poisson's equation solution at that

boundary.

The drain and source contacts are treated as particle reservoirs, and the carriers may enter or exit the device through these terminals. Regarding Poisson's equation, the contacts are considered to be Dirichlet boundaries where the bias voltages are applied. If a carrier crosses the device border in a contact, it is accounted in the calculation of the current. The metal gate contact has the potential $\phi(i, j, k)$ calculated as:

$$\phi(i, j, k) = \frac{V_g}{V_t} + dE_c + \frac{X_{sc}}{V_t} - \frac{W_m}{V_t} \quad (10)$$

in which W_m is the metal work function, X_{sc} is the silicon electron affinity and $V_t = kT/q$ is the thermal voltage.

D. Metal Grain Granularity

Since metal grain granularity (MGG) is a relevant source of variability in FinFETs [9,10,11], it is mandatory to include its impact on device fluctuation.

The MGG modelling approach which was used in the code is the one presented in [12]. First, the grains and their characteristics must be defined. The average grain size (A_g) is defined according to the metal gate material, ranging from 5 to 20 nm. The expected number of grains ($E[Ng]$) is calculated for each metal gate plane as the ratio between the surface area (SA) and the average grain size, i.e., $E[Ng] = SA/A_g$. The number of grains is drawn from a Poisson distribution. The position of each grain is randomly selected from a uniform distribution in the 2D plane of the metal gate. For each grain, the WF is randomly selected according to the WF values and their occurrence probabilities depending on the metal gate material. Lastly, the grains must be assigned to the grid, in order to define the WF value on each mesh cell. For each grid node in the gate, the distance to each grain in the plane is calculated, and the WF value from the closer grain is assigned to the node. The process is repeated for each metal gate surface.

E. Dopants Distribution

The doping atoms are randomly placed in the appropriate regions of the device. The mean number of dopants that belong to a certain region is calculated by the corresponding volume and the doping density. The actual number of doping atoms in a region is randomly picked from the Poisson distribution, the mean of which being the number calculated from the volume and doping density.

F. Carrier Initialization

The initial number of carriers and their location is not arbitrary, being defined so that initial charge neutrality is assured in the device by balancing the dopants ions. Similar to the bulk Monte Carlo method, the energy and the momentum of each electron is obtained assuming equilibrium Boltzmann distribution function.

G. Monte Carlo Routine

The Monte Carlo method, as discussed in Chapter 2, consists of a free-flight and scattering loop. The difference is that the semiconductor is no longer physically unlimited, so that the device boundaries must be taken into consideration. The

boundaries must be checked at each time step so that no carrier enters the oxide or leaves the devices through its borders.

H. Particle-Mesh Coupling

While Poisson's Equation is solved on discrete grid points, the Ensemble Monte Carlo method evaluates particle motion in continuous space coordinates. Hence, a particle-mesh (PM) coupling method is needed for both particle assignment to a mesh point and the force interpolation. The PM coupling is compound of the assignment of the particle to the mesh, Poisson's equation solution on the mesh, calculation of the mesh-defined forces and interpolation to define the forces on the particle.

There are three particle-mesh coupling methods: nearest-grid-point (NGP), nearest-element-center (NEC) and cloud-in-cell (CIC). In this work, the CIC method was used, since it is the most precise of all.

I. Particle-particle-particle-mesh (P³M) Algorithm

To properly simulate particle dynamics, all relevant scattering mechanisms must be taken into account, including carrier-carrier and carrier-ion interactions. The scattering rates used in the Monte Carlo algorithm are typically incorporated in the k-space portion of the Monte Carlo transport kernel; in the past, it was common to also model Coulomb interaction as additional scattering mechanism using the Brooks-Herring approach. This approach leads to double counting of the Coulomb force, disregards multi-ion contributions to the scattering potential and of dynamical perturbations to the Coulomb fields resulting from carriers' movement. The k-space approach for the Coulomb interaction also requires frequent evaluation of the distribution function, which is very time consuming task [13]. To avoid these difficulties, particle-particle-particle-mesh (P³M) was adopted [14]. Other methodologies, such as the fast multipole method [15] and the corrected Coulomb forces approach [16] were also used. In this work we adopt the P³M method.

The particle-particle-particle-mesh algorithm is a hybrid of two models: the particle-particle (PP) model, that uses the force law to obtain the forces of interaction and the equations of motion to obtain the state of the system (particle positions and velocities) for a certain time, and the particle-mesh (PM) model, that treats the force as a field quantity approximated on a mesh. In the PM model, the charge is assigned to a mesh, and the Poisson's equation is then solved, followed by the computation of the mesh-defined potential and its interpolation to the actual particle positions. While the PP model is suitable for small systems with long-range forces or for large systems with only a few interparticle distances that result in nonzero forces of interaction, the PM method can have a significant reduction on computational costs, at the expense of loss of resolution in the potential and force fields, so that it is appropriate only for smoothly varying forces. The P³M method combines the advantages of the PP and PM models, allowing the simulation of large correlated systems with long-range (such as Coulomb) forces [14].

The basic idea of the P³M method is to split the interparticle forces into two components: a short-range part, which is nonzero only for particles separations smaller than a certain

cutoff radius and is computed by direct particle-particle pair force summation, and a long-range part, that varies smoothly and is calculated by the particle-mesh method. Eq. (11) shows that the total force on particle i can be split in the direct forces of particles j on particle i in the short-range domain (SRD), corresponding to the first sum, and in the mesh forces of particle j on particle i in the global problem domain (GD), including the effect of material boundaries and boundary conditions on particle i , corresponding to the second sum [13].

$$\mathbf{F}_i = \sum_{j \neq i}^{\text{SRD}} \mathbf{F}_{ij}^{sr} + \sum_{j \neq i}^{\text{GD}} \mathbf{F}_{ij}^m \quad (11)$$

In Eq. (11), \mathbf{F}_{ij}^{sr} is the short-range Coulomb force of particle j on particle i , and \mathbf{F}_{ij}^m is the long-range mesh force of particles j on particle i .

The short-range Coulomb force can be written as

$$\mathbf{F}_{ij}^{sr} = \mathbf{F}_{ij}^{coul} - \mathbf{R}_{ij} \quad (12)$$

where \mathbf{F}_{ij}^{coul} is the force of particle j on particle i , given by Coulomb's law,

$$\mathbf{F}_{ij}^{coul} = \frac{q_i q_j (\mathbf{r}_i - \mathbf{r}_j)}{4\pi\epsilon |\mathbf{r}_i - \mathbf{r}_j|^3}, \quad (13)$$

where q_i and q_j are particles charges and \mathbf{r}_i and \mathbf{r}_j are particle positions. The variable \mathbf{R}_{ij} is a reference force that avoids double counting of the forces of the SRD and GD, being equal to the mesh force within the short-range domain and to the Coulomb force outside the short-range domain. The Coulomb force is kept restricted to the SRD, where the reference force goes smoothly to zero, and outside the short-range domain cutoff radius (\mathbf{r}_{sr}), the force is kept equal to the mesh forces, so that the reference force corresponds to the point particle force law.

To properly integrate P³M with the EMC method to account for carrier-carrier and carrier-impurity interactions for uniform meshes, Wordelman and Ravaioli [17] followed Hockney's approach and obtained the reference forces used in this work, shown in Eq. (14).

$$\mathbf{R}_{ij}(r) = \frac{q_i q_j}{4\pi\epsilon} \cdot \left\{ \begin{array}{l} \frac{1}{35r_{sr}^2} \left(\begin{array}{l} 224\xi - 224\xi^3 + 70\xi^4 + 48\xi^5 - \\ -21\xi^6 \end{array} \right), \\ \text{for } 0 \leq r \leq r_{sr}/2 \\ \frac{1}{35r_{sr}^2} \left(\begin{array}{l} \frac{12}{\xi^2} - 224 + 896\xi - 840\xi^2 + \\ 224\xi^3 + 70\xi^4 - 48\xi^5 + 7\xi^6 \end{array} \right), \\ \text{for } r_{sr}/2 \leq r \leq r_{sr} \\ \frac{1}{r^2}, \text{ for } r > r_{sr} \end{array} \right\} \quad (14)$$

where r is the interaction radius and $\xi = 2r/r_{sr}$.

J. Poisson's Equation Solver

Monte Carlo device simulation requires the long-range force that is calculated from the potential profile, which, in turn, is obtained as a solution of Poisson equation, given by

$$\nabla \cdot (\epsilon_s \nabla \psi) = -\rho \quad (15)$$

where ψ is the electrical potential, ρ is the free charge density and ϵ_s is the material permittivity. The 3-D Poisson equation is solved using the strongly-implicit method [18].

V. BULK SIMULATOR RESULTS

Prior to attempting device simulation, bulk simulations were performed to verify proper modelling of basic silicon properties, such as average drift velocity and average carrier energy, which are later implicitly used to simulate FinFETs. The relation between the velocity and the electric field is shown in Fig. 6, and the dependence of energy with the electric field is shown in Fig. 7. The carriers' velocity and energy are in good agreement with the experimental data of Canalli et al. [19] and the simulation results of Jacoboni and co-workers [20].

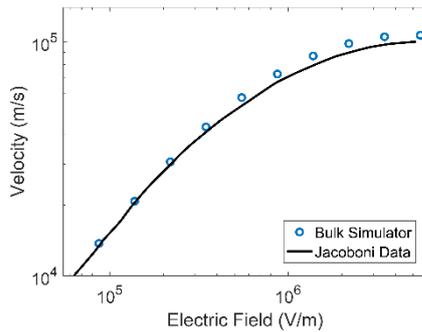


Fig. 6. Bulk silicon velocity dependence upon the electric field.

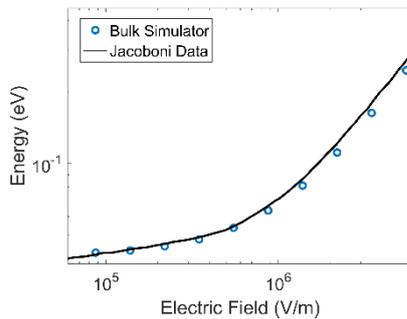


Fig. 7. Average carrier energy in bulk silicon as a function of the electric field.

VI. MONTE CARLO DEVICE SIMULATOR RESULTS

Some relevant simulation results that validate the 3D FinFET simulator are presented next. A n-FinFET with channel length $L_{ch} = 18$ nm, $W_{fin} = 8$ nm and $H_{fin} = 42$ nm was simulated. The total device length (L) is 78 nm, which includes the source (0 – 30 nm), the channel (30 nm – 48 nm) and the drain (48 nm – 78 nm). The total device depth (D) is 57.2 nm. The total device width (W) is 10.4 nm, which includes the thickness of the gate dielectrics located in the interface with the side channels. The simulation time was 15 ps. The simulations were performed with $V_G = V_D = 0.9$ V and $V_S = V_B = 0$ V, except when said differently. TiN was used as metal gate material, which has average grain size of 20 nm and work function of $WF = 4.6$ eV with 60% probability and $WF = 4.4$ eV with 40% probability [21]. Surface roughness scattering is treated as being 15% diffusive and 85% specular [22]. Fig.

8 shows the basic, simplified structure of the FinFET devices which were implemented in the simulator.

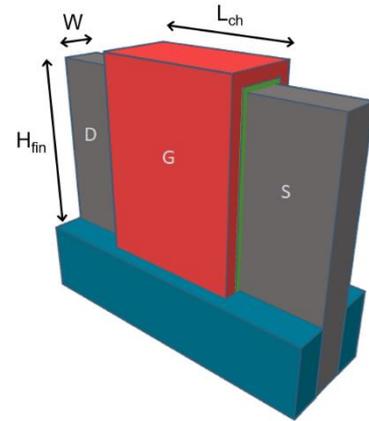


Fig. 8. Simplified schematic drawing of the basic FinFET structure. The silicon area is depicted in gray, the metal gate is presented in red, the silicon oxide is depicted in blue and the gate dielectric is shown in green.

The device length is in the direction from source to drain, i.e., along the channel length (defined as L-direction). The device depth is defined as the oxide-substrate direction (D-direction). The width of the device is aligned with the channel width (W-direction).

The random placement of the dopants in a device is shown in Fig. 8 for acceptors and donors doping concentration of $N_A = 5 \cdot 10^{18} \text{ cm}^{-3}$ and $N_D = 10^{19} \text{ cm}^{-3}$ respectively.

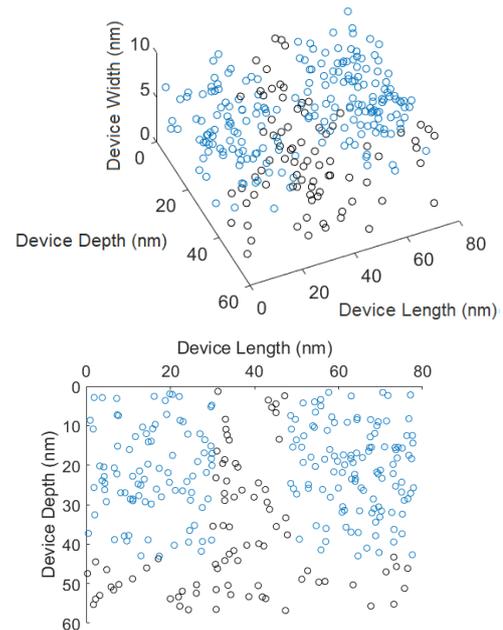


Fig. 9. Dopant atoms randomly distributed in the device. The acceptor atoms are represented with black circles, and the donors with blue circles.

Metal grain granularity is depicted in Fig. 9, which presents the electric potential on a side gate. Fig. 9 clearly shows the presence of metal grains and illustrates the existence of regions with different work functions that lead to fluctuations on the gate potential.

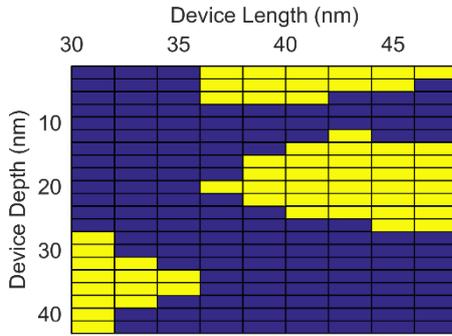


Fig. 10. Electric potential distribution over one of the side gates of a FinFET, showing the presence of metal grains with different characteristics. The yellow cells (1.103 V) have a potential 0.2 V higher than the blue ones (0.903 V), due to the difference in the work function in those regions.

Fig. 10 (a) and (b) shows respectively the carriers' average velocity and energy. It is possible to see that the energy peak is located near the drain end, as expected, and experimentally verified.

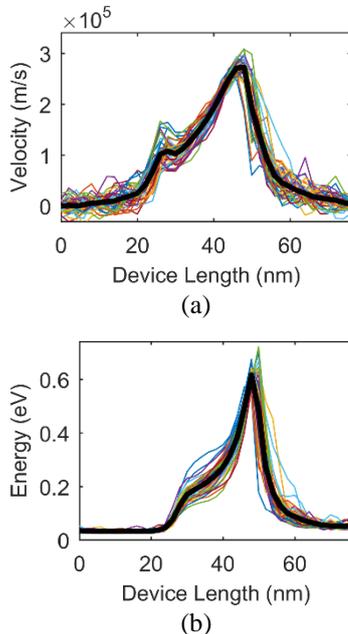


Fig. 11. Velocity (a) and energy (b) of the ensemble of carriers in different devices. The black line represents the average values for a set of devices.

The electric potential was plotted for different cut-planes in the length and the depth of the device, with the width direction fixed as the center of the fin, i.e., equal to $W/2$. In Fig. 11, we show the simulation results for the electrostatic potential with $V_D = V_S = V_B = 0$ V. In Fig. 12, the same plot is shown, but with the drain bias of $V_D = 0.9$ V and $V_S = V_B = 0$ V. The total simulation time is 15 ps. It is possible to see that the donor dopants are randomly distributed in the source and in the drain region and they significantly perturbate the potential. In the channel region ($L = 30$ nm – 48 nm, $H = 1.2$ nm – 43.2 nm and $W = 1.2$ nm – 9.2 nm), there are also significant variations of the electrostatic potential caused by the substrate dopants.

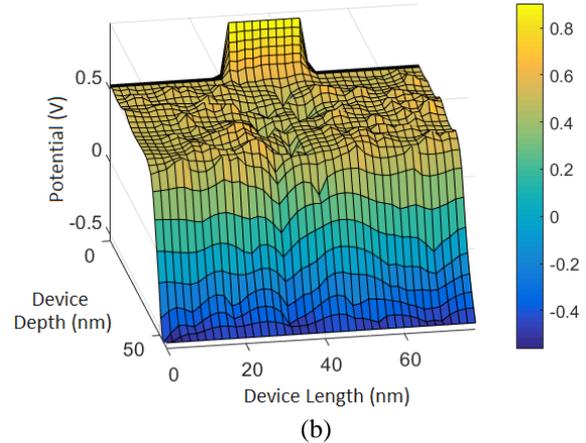
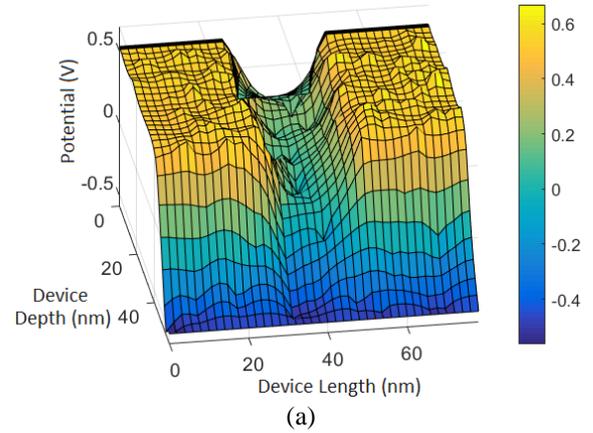
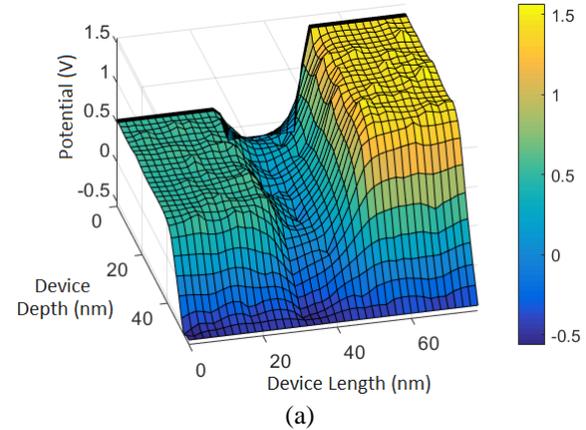


Fig. 12. Electrostatic potential along the device length and depth with $V_D = V_S = V_B = 0$ V and (a) $V_G = 0.1$ V, and (b) $V_G = 0.9$ V.

With $V_D = V_S = V_B = 0$ V, the source and drain contacts are at the same electrostatic potential, and the gate bias impact is clearly seen from Figs. 11 (a) and (b). The case when voltage of 0.9 V is applied on the drain is shown in Fig. 12.



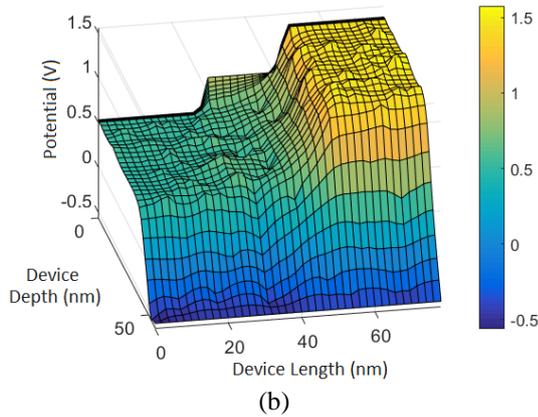


Fig. 13. Electrostatic potential in the device as a function of its length and depth with $V_D = 0.9$ V, $V_S = V_B = 0$ V, and (a) $V_G = 0.1$ V, and (b) $V_G = 0.9$ V.

The electrostatic potential was also plotted as a function of the position in the direction of device width and length, with the position in the depth direction fixed as the top channel silicon/gate dielectric interface. Fig. 13 displays the results for $V_G = 0.9$ V, $V_D = V_S = V_B = 0$ V, and in Fig. 14 we show the results obtained for $V_G = V_D = 0.9$ V and $V_S = V_B = 0$ V after the 15 ps simulation.

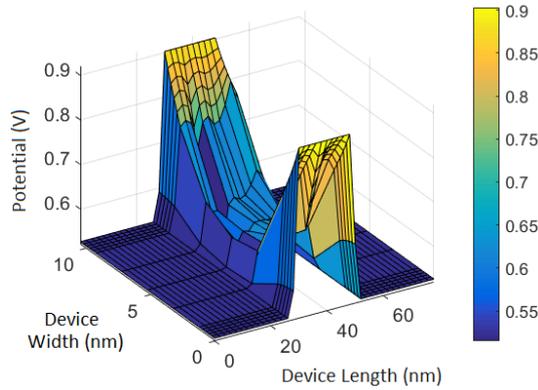


Fig. 14. Electrostatic potential along device length and width for equilibrium condition with $V_G = 0.9$ V, $V_D = V_S = V_B = 0$ V.

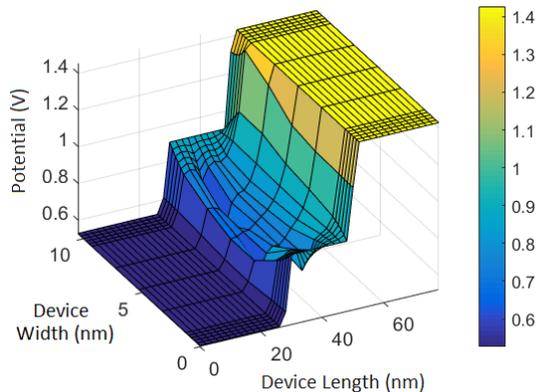


Fig. 15. Electrostatic potential in the device as a function of its width and depth with $V_G = V_D = 0.9$ V and $V_S = V_B = 0$ V.

Fig. 14 shows electron density distribution of a device along the directions of its depth and length after the drain voltage of 0.9 V is applied and the device is simulated for 15 ps.

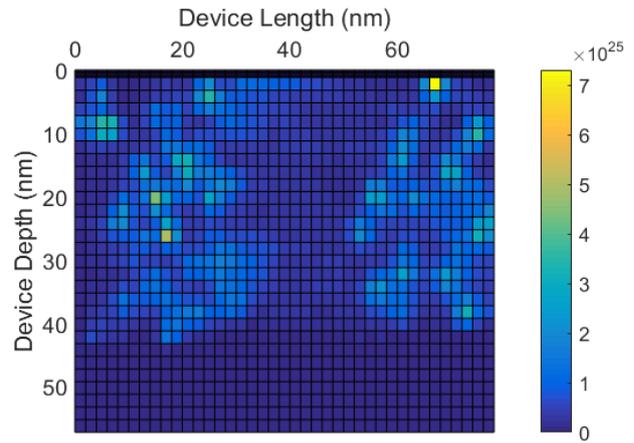


Fig. 16. Electron density [m^{-3}] along device length and depth after 15ps simulation for a slice in the middle of the channel width ($W/2$).

Note that, since the plot in Fig. 15 represents a 2-D slice taken from the middle of the fin width, the top channel carriers are captured in the figure as a high concentration near the oxide-silicon interface with depth of 1.2 nm.

Fig. 16 shows carrier density also in the directions of device length and depth and for the same simulation conditions, but with the position on the width direction defined as the interface between one of the side channels and the gate dielectric. It is possible, therefore, to capture the behavior of the side channel as well. In both Figs. 15 and 16, pinch-off effect is clearly seen near the drain end (around 48 nm in the device length direction), where the electron density becomes very low.

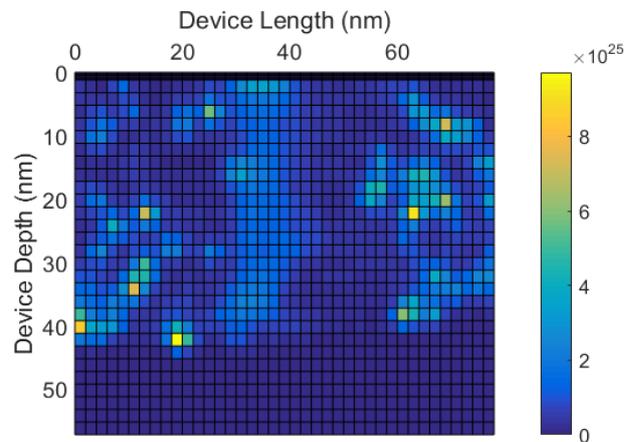


Fig. 17. Electron density [m^{-3}] along device length and depth after 15ps simulation for a slice that captures one of the side channels.

The behavior of the drain current of the n-FinFET is shown in Fig. 17 as the $I_D - V_G$ curve, and in Fig. 18 as $I_D - V_D$ curve. These results correspond to well established values for MOS transistors.

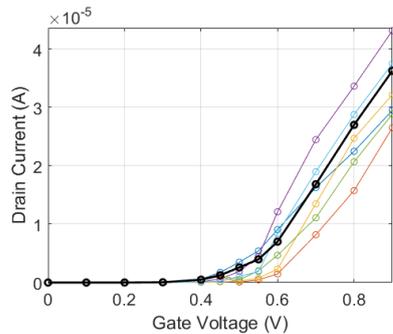


Fig. 18. $I_D - V_G$ curve for different devices, depicting the variability between devices with the same parameters.

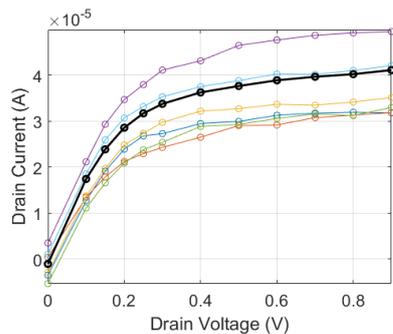


Fig. 19. $I_D - V_D$ curve for different devices with the same parameters as in Fig. 17.

VII. CONCLUSION

This work presented a 3-D TCAD particle-based semiconductor device simulator. First, it was discussed that semiconductor device simulation requires self-consistent solution of the transport and field equations. While the particle dynamics was calculated in this work by solving BTE using Monte Carlo technique, the electric fields were obtained by numerically solving Poisson's equation. It was addressed that device simulation is basically compound of the initialization of material parameters, device structure and carrier distribution, Monte Carlo random walk/scattering, solution of Poisson's equation and interpolation of force to particle location. During this process, the boundaries and the contacts are constantly checked, so that no carrier leaves the device unaccounted for.

The n-FinFET 3D device simulator that was implemented provides established results for the device transfer and output characteristics. The behavior of electron's velocity and energy along the channel matches the expected results, with both quantities reaching their peak values near the drain end. The drain current versus gate voltage and drain current versus drain voltage curves also exhibited the usual n-FinFET characteristics. The electrostatic potential was plotted for different cross-sections in the directions of device length, width and depth. Both the electrostatic potentials and the carrier densities showed good agreement with the expected tendencies, corroborating that the 3-D particle-based device simulator presented in this work is adequate to simulate modern FinFETs.

ACKNOWLEDGEMENTS

This work was partially supported by Conselho Nacional

de Desenvolvimento Científico e Tecnológico – CNPq. One of the authors (Dragica Vasileska) would like to acknowledge the financial support from the National Science Foundation (USA) under contract number ECCS 1542160.

REFERENCES

- [1] S. Saxena et al., "Variation in Transistor Performance and Leakage in Nanometer-Scale Technologies," in *IEEE Transactions on Electron Devices*, vol. 55, no. 1, pp. 131-144, Jan. 2008, doi: 10.1109/TED.2007.911351.
- [2] C. Shin, X. Sun and T. K. Liu, "Study of Random-Dopant-Fluctuation (RDF) Effects for the Trigate Bulk MOSFET," in *IEEE Transactions on Electron Devices*, vol. 56, no. 7, pp. 1538-1542, July 2009, doi: 10.1109/TED.2009.2020321.
- [3] S. K. Saha, "Modeling Process Variability in Scaled CMOS Technology," in *IEEE Design & Test of Computers*, vol. 27, no. 2, pp. 8-16, March-April 2010, doi: 10.1109/MDT.2010.50.
- [4] M. Fan, V. P. Hu, Y. Chen, P. Su and C. Chuang, "Analysis of Single-Trap-Induced Random Telegraph Noise on FinFET Devices, 6T SRAM Cell, and Logic Circuits," in *IEEE Transactions on Electron Devices*, vol. 59, no. 8, pp. 2227-2234, Aug. 2012, doi: 10.1109/TED.2012.2200686.
- [5] T. Nagumo, K. Takeuchi, T. Hase and Y. Hayashi, "Statistical characterization of trap position, energy, amplitude and time constants by RTN measurement of multiple individual traps," 2010 International Electron Devices Meeting, 2010, pp. 28.3.1-28.3.4, doi: 10.1109/IEDM.2010.5703437.
- [6] M. Costato and L. Reggiani, "Nonparabolicity and overlap effects on transport problems in the Froehlich and Paranjape approach." *J. Phys. C: Solid State Phys.*, vol. 5, no. 2, pp. 159, Jan. 1972.
- [7] K. Tomizawa. *Numerical Simulation of Submicron Semiconductor Devices*. Boston, MA, USA: Artech House, 1993.
- [8] D. Vasileska; S. M. Goodnick, G. Klimeck. *Computational Electronics: semiclassical and quantum device modeling and simulation*. CRC press, 2016.
- [9] X. Wang, A. R. Brown, Binjie Cheng and A. Asenov, "Statistical variability and reliability in nanoscale FinFETs," 2011 International Electron Devices Meeting, Washington, DC, USA, 2011, pp. 5.4.1-5.4.4, doi: 10.1109/IEDM.2011.6131494.
- [10] D. Nagy, G. Indalecio, A. J. García-Loureiro, M. A. Elmessary, K. Kalna and N. Seoane, "Metal Grain Granularity Study on a Gate-All-Around Nanowire FET," in *IEEE Transactions on Electron Devices*, vol. 64, no. 12, pp. 5263-5269, Dec. 2017, doi: 10.1109/TED.2017.2764544.
- [11] A. Sudarsanan, S. Venkateswarlu and K. Nayak, "Impact of Fin Line Edge Roughness and Metal Gate Granularity on Variability of 10-nm Node SOI n-FinFET," in *IEEE Transactions on Electron Devices*, vol. 66, no. 11, pp. 4646-4652, Nov. 2019, doi: 10.1109/TED.2019.2941896.
- [12] P. H. Vardhan, S. Mittal, S. Ganguly and U. Ganguly, "Analytical Estimation of Threshold Voltage Variability by Metal Gate Granularity in FinFET," in *IEEE Trans on Electron Devices*, vol. 64, no. 8, pp. 3071-3076, Aug. 2017, doi: 10.1109/TED.2017.2712763.
- [13] S. S. Ahmed, "Modeling quantum and coulomb effects in nanoscale devices", Ph.D. dissertation, Arizona State Univ., Tempe, AZ, USA, 2005.
- [14] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles*. CRC Press, 1988.
- [15] L. Greengard, and V. Rokhlin. "A fast algorithm for particle simulations." *Journal of computational physics*, vol. 135, no. 2, pp. 280-292, 1997, doi: 10.1006/jcph.1997.5706.
- [16] W. J. Gross, D. Vasileska and D. K. Ferry, "A novel approach for introducing the electron-electron and electron-impurity interactions in particle-based simulations," in *IEEE Electron Device Letters*, vol. 20, no. 9, pp. 463-465, Sept. 1999, doi: 10.1109/55.784453.

- [17] C. J. Wordelman and U. Ravaioli, "Integration of a particle-particle-particle-mesh algorithm with the ensemble Monte Carlo method for the simulation of ultra-small semiconductor devices," in *IEEE Transactions on Electron Devices*, vol. 47, no. 2, pp. 410-416, Feb. 2000, doi: 10.1109/16.822288.
- [18] H. L. Stone, "Iterative solution of implicit approximations of multidimensional partial differential equations", *SIAM Journal on Numerical Analysis*, vol. 5, no. 3, pp. 530-558, 1968, doi: 10.1137/0705044.
- [19] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alberigi-Quaranta, "Electron drift velocity in silicon", *Phys. Rev. B*, vol. 12, no. 6, 1975, doi: 10.1103/PhysRevB.12.2265.
- [20] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, "A review of some charge transport properties of silicon", *Solid-State Electronics*, vol. 20, no. 2, pp. 77-89, 1977, doi: 10.1016/0038-1101(77)90054-5.
- [21] H. F. Dadgour, K. Endo, V. K. De and K. Banerjee, "Grain-Orientation Induced Work Function Variation in Nanoscale Metal-Gate Transistors—Part I: Modeling, Analysis, and Experimental Validation," in *IEEE Transactions on Electron Devices*, vol. 57, no. 10, pp. 2504-2514, Oct. 2010, doi: 10.1109/TED.2010.2063191.
- [22] F.M. Bufler, L. Smith, "3D Monte Carlo simulation of FinFET and FDSOI devices with accurate quantum correction," *J Comput Electron*, 12, 651–657, 2013. <https://doi.org/10.1007/s10825-013-0518-z>