Fermi-level effects in semiconductor processing: A modeling scheme for atomistic kinetic Monte Carlo simulators

I. Martin-Bragado[a)]
Department of Electronics, University of Valladolid, Escuela Técnica Superior de Ingenieros de Telecomunicacion (ETSIT) Campus Miguel Delibes, 47011 Valladolid, Spain and Synopsys, Gesellschaft mit Beschränkter Haftung Karl-Hammerschmidth-Strasse 34, D-85609 Aschheim/Dornach, Germany

P. Castrillo, M. Jaraiz, R. Pinacho, J. E. Rubio, and J. Barbolla
Department of Electronics, University of Valladolid, Escuela Técnica Superior de Ingenieros de Telecomunicacion (ETSIT) Campus Miguel Delibes, 47011 Valladolid, Spain

V. Moroz
Synopsys, Inc., 700 Middlefield Road, Mountain View, California 94043

(Received 23 November 2004; accepted 1 August 2005; published online 12 September 2005)

Atomistic process simulation is expected to play an important role for the development of next generations of integrated circuits. This work describes an approach for modeling electric charge effects in a three-dimensional atomistic kinetic Monte Carlo process simulator. The proposed model has been applied to the diffusion of electrically active boron and arsenic atoms in silicon. Several key aspects of the underlying physical mechanisms are discussed: (i) the use of the local Debye length to smooth out the atomistic point-charge distribution, (ii) algorithms to correctly update the charge state in a physically accurate and computationally efficient way, and (iii) an efficient implementation of the drift of charged particles in an electric field. High-concentration effects such as band-gap narrowing and degenerate statistics are also taken into account. The efficiency, accuracy, and relevance of the model are discussed. © *2005 American Institute of Physics*. [DOI: 10.1063/1.2041847]

I. INTRODUCTION

For a device with a 20-nm physical channel length, which is expected to be in production by 2009,[1] the discrete nature of dopant atoms is likely to have a significant impact on the device characteristics. The International Technology Roadmap for Semiconductors cites "atomistic process modeling" as one of the long-term (beyond 2010) modeling and simulation requirements for which "manufacturable solutions are not known."[1] Therefore, physically based models that take into account the discrete properties of each defect should be developed to be included in atomistic simulators as a way to accurately predict the behavior of such current and future devices. With each technology generation, the junctions are required to become shallower, but without a significant increase of the series resistance. This is satisfied by employing high doping levels which introduce a variety of effects related to the high concentration and spatial variations of electrical charges.

Among the atomistic simulation techniques, the nonlattice kinetic Monte Carlo (kMC) approach is, probably, the best choice for the size and time scales involved in the fabrication of a typical semiconductor device.[2–5] This technique has already proven to produce accurate results for different process simulation scenarios including transient-enhanced diffusion (TED),[2,6–8] dopant activation/deactivation phenomena,[7–11] and damage accumulation.[12,13] Coupled to a device simulator, it was used to describe the effect of statistical variations in field-effect transistors.[14] There have been some reports including charge effects in both lattice[15] and nonlattice kMC simulators[3,4,16] but without any details on how the charge effects are modeled. The impact of electrical charges on dopant and defect diffusion in silicon is described in terms of continuum physics in Ref. 17 and it has recently been reformulated within the atomistic kMC framework in Ref. 16. Such theoretical derivations can be done in a rather systematic way. However, their implementation in an atomistic process simulator poses a number of practical difficulties that arise from the inherent discreteness of the atomistic simulation.

In this work we describe a three-dimensional charge model for point defects and dopants as implemented in the atomistic nonlattice kMC simulator DADOS.[3,18] We focus on specific issues associated with the atomistic nature of the kMC approach and propose practical solutions to these issues. In Sec. II we summarize the guidelines of the physical model (for more details see Ref. 16). The details of the implementation algorithm are described in Sec. III. Finally, some illustrative examples are presented and discussed in Sec. IV.

II. THE PHYSICAL MODEL

In the conventional continuum models each defect- or dopant-species-type is represented as a unique concentration variable with an effective diffusivity that depends on the local Fermi level (e_F).[19] All possible charge states are lumped into that single concentration variable and are used to determine the effective diffusivity of the species.

[a)]Electronic mail: ignacio.martin-bragado@synopsys.com

 © 2005 American Institute of Physics

Downloaded 22 Sep 2005 to 157.88.111.70. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

TABLE I. Energy levels for the charged states of silicon self-interstitial (I^j), vacancy (V^j), interstitial-boron pair (B_I^j), interstitial-arsenic pair (As_I^j), and arsenic-vacancy pair (AsV^j). Only independent parameters are listed in the table. The values for charge levels $e(j+1,j)$ are given for $T=0$ K with no band-gap narrowing and they are measured from the valence-band edge.

	I^-	I^0	I^+	V^{--}	V^-	V^0	V^+	V^{++}	B_I^-	B_I^0	B_I^+	As_I^0	As_I^+	AsV^-	AsV^0	AsV^+
$e(j+1,j)$ (eV)	1	0.4		1.06	0.6	0.03	0.13		0.8	1.04		0.1		0.77	0.3	
$D_{m,0}\times 10^{-3}$ (cm^2/s)	5	5	5	1	1	1	1	1	1	1	1	5	5	1	1	1
E_m (eV)	1	1	1	0.4	0.4	0.4	0.6	0.8	0.36	0.2	1.1	1.35	1.35	1.7	1.4	1.4
$D_{bk,0}\times 10^{-3}$ (cm^2/s)									400				5			1
E_b (eV)									0.1				0.1			1
$[X^0]_0^*(\times 10^{25}$ cm$^{-3})$		690				2.6										
E_f (eV)		3.85				3.7										

In our atomistic model, each individual particle is described by its position, type, and charge state. The particle types are defects (interstitials or vacancies, I or V), dopants (B, As,...), and dopant-defect pairs ($B_I, As_I, AsV, \ldots$). The ratio of the relative concentrations of a particle X in the charge states j and $j+1$ is given by

$$\frac{[X^j]}{[X^{j+1}]} = \exp\left[\frac{e_F - e(j+1,j)}{k_B T}\right], \qquad (1)$$

where $e(j+1,j)$ is the energy level associated with the charge transition, k_B the Boltzmann constant, and T the temperature. From an atomistic point of view, Eq. (1) can be interpreted as the probability ratio for a particle X to be in the charge states j or $j+1$. Charge transitions are assumed to be much faster than atomic transport[20] and, therefore, charge state probability ratios follow Eq. (1) instantaneously. We assume that $e(j+1,j)$ scales with the temperature in the same way as the band-gap energy $E_g(T)$.[21] The values used for the energy levels $e(j+1,j)$ at $T=0$ are listed in Table I. The experimental data and theoretical calculations used to obtain these numbers are described in Refs. 11 and 16. Substitutional shallow dopants are assumed to be always ionized $(B^-, As^+, \ldots)$ at annealing temperatures.

The concentrations of neutral defects in equilibrium conditions ($[I^0]^*$ and $[V^0]^*$) are controlled by the balance between the defect emission and capture at the surfaces. These concentrations are independent of e_F and have activation energies equal to the corresponding formation energies $[E_f(I^0)$ and $E_f(V^0)]$. The values for these formation energies, as well as the prefactors for the equilibrium concentrations ($[I^0]_0^*$ and $[V^0]_0^*$), are shown in Table I. The concentrations of charged I's and V's are controlled by charge transitions [Eq. (1)] and are, in turn, e_F dependent.[16] Therefore, the formation energies of charged defects can be obtained from $E_f(X^0)$ and e_F using:

$$E_f(X^{j+1}) = E_f(X^j) + e_F - e(j+1,j). \qquad (2)$$

In the absence of an electric field, the particles jump randomly with a frequency

$$\nu_m = \frac{6D_{m,0}}{\lambda^2}\exp\left(-\frac{E_m}{k_B T}\right),$$

$D_{m,0}$ being a diffusivity prefactor, λ the jump distance, and E_m the migration energy.[3] We take λ equal to the second neighbor's distance in silicon ($\lambda=0.384$ nm). Each charge state has an independent jump frequency that does not depend on e_F. The values that we use for $D_{m,0}$ and E_m can be found in Table I. Substitutional dopants are assumed to be immobile. Therefore, dopant diffusion is determined by the concentration and diffusivity of mobile defect-dopant pairs.

An electric field ξ introduces a gradient of the formation energy of a charged particle X^j, inducing an asymmetry in its jump probabilities:[3,16]

$$\frac{\nu_m^+}{\nu_m^-} = \exp\left(\frac{jq\lambda\xi}{k_B T}\right), \qquad (3)$$

ν_m^+ and ν_m^- being the jump frequencies in the positive and negative directions, respectively, ξ the electric-field component, and q the absolute value of the electron charge. This asymmetry is the microscopic origin of the macroscopic drift of the charged particles in an electric field.

Interaction between particles can occur within a given interaction distance. For example, if a jumping point defect X^j gets near a dopant A^k (with $k=+1$ for donors and $k=-1$ for acceptors), a dopant-defect pair AX can be formed:

$$X^j + A^k \leftrightarrow AX^{j+k}. \qquad (4)$$

Reactions between particles with a charge of the same sign are forbidden due to the short-range electrostatic repulsion. Long-range electrostatic attractions or repulsions are accounted for by the drift of the charged particles in the electric field.

The reverse process (i.e., the breakup of an AX^j pair) will happen with a rate:

$$\nu_{bk}(AX^j) = \frac{6D_{bk,0}(AX^j)}{\lambda^2}\exp\left[-\frac{E_b(AX^j) + E_m(X^{j-k})}{k_B T}\right],$$

where $D_{bk,0}$ is the breakup prefactor in diffusivity units and E_b is the binding energy. Note that the reactions considered in the expression (4) conserve the charge. Other pairing or breakup reactions could be seen as a combination of a charge-conserving reaction and a charge transition [included in Eq. (1)]. Breakup prefactors and binding energies for the different charge states of a given AX pair are linked by microscopic reversibility considerations as follows:[16,17]

$$E_b(AX^{j+1}) = E_b(AX^j) + e_{AX}(j+1,j) - e_X(j+1-k, j-k), \qquad (5)$$

Downloaded 22 Sep 2005 to 157.88.111.70. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

$$D_{bk,0}(AX^{j+1}) = D_{bk,0}(AX^{j}) \frac{D_{m,0}(X^{j+1-k})}{D_{m,0}(X^{j-k})}. \tag{6}$$

The concentration of AX^j in local equilibrium is established by the balance of pairing and breakup rates:

$$[AX^j] = \upsilon_{\text{capt}}[A^k][X^{j-k}] \frac{\nu_m(X^{j-k})}{\nu_{bk}(AX^j)}, \tag{7}$$

where υ_{capt} is the effective capture volume. In our case, $\upsilon_{\text{capt}}=2\lambda^3$. The activation energy for the equilibrium concentration of pairs resulting from pairing reactions between neutral defects and charged substitutional dopants ($A^k + X^0 \leftrightarrow AX^k$) are independent of e_F and we use the corresponding binding energies as reference points to compute the E_b for the other charge states. The independent values $D_{bk,0}$ and E_b that we take for B_I, As_I, and AsV are shown in Table I.

The model obtains the effective dopant diffusivity as a natural consequence of having the right proportion of the different charged particles, based on the local e_F.[16] The relationship between the microscopic parameters described here and the parameters used in continuum simulators can be found in Ref. 16.

III. MODEL IMPLEMENTATION

Once the theoretical basis and key assumptions of the model are introduced, we are going to deal with the implementation issues associated with the atomistic framework. In particular, in order to perform a realistic simulation of impurity diffusion in silicon, it is necessary to know:

- How to compute the Fermi level from a distribution of discrete charged particles.
- How to update the charge for the particles with several possible charge states, in order to follow the space-dependent relative probabilities of Eq. (1).
- How to implement drift of the charged particles in an electric field.

A. Computing the charge distribution

Electron concentration is treated in a continuum way. To compute the electron concentration, and subsequently e_F, the simulation domain is divided into a tensor product grid with uniform spacing where each grid cell is a rectangular box. The boxes in a grid have the same size and are small enough (≈ 1 nm^3) to assume that the electron concentration is constant in each element (box) of the grid.

To calculate the electron concentration, our algorithm does not only count the number of substitutional dopants (assumed to be always ionized) inside each box, but also averages the electron concentration over neighboring boxes and accounts for the band-gap narrowing as described below. Once the electron concentration is calculated, it is used to update e_F in each box, using Fermi-Dirac statistics and the charge neutrality assumption. Unless the smoothing algorithm is used, the discrete, atomistic charge distribution leads to a computed electron concentration that is artificially high (or low) near the position of active dopants; and this concentration would depend on the size of the discretization boxes used in the simulator, changing abruptly from one box to another, and generating unphysical artifacts in the simulation, as we will see in Sec. IV.

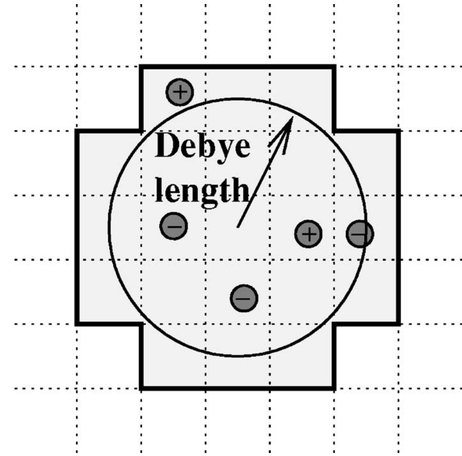

FIG. 1. Illustration of the charge smoothing. The electron concentration in each box of the grid is computed as the average electron concentration in the neighborhood of the box, where the neighborhood radius is the Debye length.

To compute the charge that is affecting the electron concentration in a particular box we use a radius r_D (see Fig. 1) that is set equal to the temperature-dependent Debye length,[21]

$$r_D = (4\pi q^2 N_{\text{dop}}/T)^{-1/2},$$

N_{dop} being the local density of active dopants. This charge averaging procedure accounts for effects such as the transition regions in *pn* junctions without the need to solve Poisson's equation. This is an intermediate approach between the local charge neutrality approximation and the exact electrostatic solution and it is good enough for the high doping levels commonly involved in deep submicron technologies.

Doping-induced energy shifts in the conduction-band minimum and valence-band maximum have been included in the model. The two components of the fundamental band-gap narrowing Δe_v and Δe_c are computed using the equations from Ref. 22 with the parameters extracted from Ref. 23. For all other electronic energy levels our model assumes that they vary in proportion with the band-gap E_g. In a region where the conduction- and valence-band edges are shifted by Δe_c and Δe_v relative to the no band-gap narrowing case where $e_c = E_g$ and $e_v = 0$, a charge level $e(j+1,j)$ will be shifted by

$$\Delta e(j+1,j) = \left[1 - \frac{e(j+1,j)}{E_g}\right]\Delta e_v + \frac{e(j+1,j)}{E_g}\Delta e_c.$$

B. Charge update scheme

As we have mentioned in Sec. II, our model assumes that the electronic transitions are much faster than the atomic diffusion and reactions. Therefore, it is necessary to implement a mechanism to update the charge distribution (and e_F) that follows the atomic jumps. The probability ratios depend

Downloaded 22 Sep 2005 to 157.88.111.70. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

on e_F [Eq. (1)], so it is necessary to update them each time it changes. The reasons for a local change of e_F can be

- Particles diffusing between boxes with different e_F.
- Change of the electron concentration in a particular box.

Furthermore, each time a new particle appears or disappears because of pairing or break-up reactions, it is necessary to ensure that the charge state of the new particle is consistent with the local e_F.

Our model implements three different mechanisms to maintain the particles in the correct e_F-dependent ratios during the simulation. These mechanisms update the electron concentration in the grid boxes, the charge state of migrating particles, and the charge state of particles involved in pairing and breakup reactions. All of them are performed at the same time, but apply to the different possible cases listed above.

1. Updating the electronic concentration and charge state ratios

The model computes the probability P^j of each species being in a particular charge state j as

$$P^j = \frac{[X^j]}{\sum_i [X^i]}, \quad j = 0, -, +, --, ++. \tag{8}$$

We need an updating method that periodically reviews all the particles and updates e_F and the proportions of charged particles in each grid box. The algorithm smooths out the charge distribution, computes e_F in the box using the charge neutrality assumption, and applies Eq. (8) to establish the appropriate charge ratios. This operation slows down the simulation, so it is crucial to perform the updates frequently enough to reflect significant changes in e_F, but without unnecessary waste of the central processing unit (CPU) time.

2. Updating mobile particles

Mobile particles see different e_F values when jumping from one box to another. Therefore, a charge update is needed at least each time a particle crosses the (artificial) boundary between the boxes. We implemented this by updating the charge state of the particles each time they jump to another box. Nevertheless, some species will diffuse faster than others, and then they will be updated more frequently. Using Eq. (8) directly to perform these updates would lead to an artifact in the simulation, because fast diffusing species will be updated more frequently, so they will have more opportunities to change their charge state than the slower ones, and then we would get an artificial increase in the slow diffusing species concentration. To prevent this situation from happening, we have modified Eq. (8) by explicitly introducing there the migration frequency. We assume that the probability for a particle that is initially in a charge state i, to switch to a charge state j after a diffusion step $P_{\text{dyn}}^{i\to j}$, does not depend on the initial state and we will simply denote it by P_{dyn}^{j}. The balance of the switch rates between the states i and j becomes

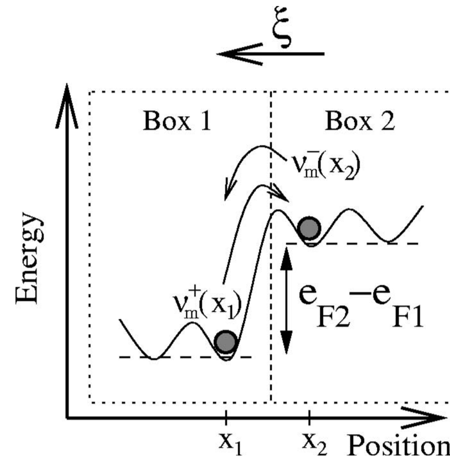

FIG. 2. A positive singly charged particle jumping from x_1 (in box 1) to x_2 (in box 2) finds an additional energy barrier which depends on the local e_F of the two boxes. For no band-gap narrowing, the barrier will be $(e_F)_2 - (e_F)_1$. It corresponds to the existence of an electric field ξ in the negative direction.

$$[X^i]\nu_m(X^i)P_{\text{dyn}}^{j} = [X^j]\nu_m(X^j)P_{\text{dyn}}^{i},$$

and therefore

$$P_{\text{dyn}}^{j} = \frac{P^j\nu_m(X^j)}{\sum_i P^i\nu_m(X^i)}.$$

3. Updating charge in pairing and breakup reactions

After pairing or breakup reactions [Eq. (4)], some species appear and disappear in a box. To ensure that the concentration of the charged species is in a correct proportion, a breakup/pairing charge update mechanism is implemented which computes the probability for new particles to be in a particular charge state j as

$$P_{\text{pair}}^{j} = \frac{P^j\nu_{bk}(X^j)}{\sum_i P^i\nu_{bk}(X^i)},$$

for pairing reactions, and

$$P_{\text{break}}^{j} = \frac{P_{\text{dyn}}^{j}\delta_j}{\sum_i P_{\text{dyn}}^{i}\delta_i},$$

for breakup reactions, where δ_i is 1 for allowed reactions and zero otherwise.

C. Drift in an electric field

Since e_F is constant inside each box, but can vary across adjacent boxes, the electric drift only applies to particles moving from one box to another. Referring to Fig. 2 and assuming $\nu_m^+(x_1)$ to be the jump frequency of a particle X^j from x_1 to x_2 (in an adjacent box) and $\nu_m^-(x_2)$ to be the jump frequency from x_2 to x_1, the ratio is

Downloaded 22 Sep 2005 to 157.88.111.70. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

$$\frac{\nu_m^+(x_1)}{\nu_m^-(x_2)} = \exp\left[\frac{(E_f)_1 - (E_f)_2}{k_B T}\right], \tag{9}$$

where E_f is the space-dependent formation energy of the charged particle X^j that can be evaluated at each position (x_1 or x_2) using Eq. (2). Therefore, the drift is implemented as an additional migration barrier $(E_f)_2-(E_f)_1$ to jump from x_1 to x_2, when $(E_f)_1<(E_f)_2$.

To visualize a simple case without the band-gap narrowing [i.e., when $e(j+1,j)_1=e(j+1,j)_2$], Eq. (9) can be simplified to

$$\frac{\nu_m^+(x_1)}{\nu_m^-(x_2)} = \exp\left[-j\frac{(e_F)_2 - (e_F)_1}{k_B T}\right], \tag{10}$$

which is the situation shown in the Fig. 2. Equation (10) is equivalent to the previously reported Eq. (3).[3,16] However, Eq. (9) is more general, since it can also account for situations where different charge states see different effective electric fields (for instance, when there is band-gap narrowing).

IV. RESULTS AND DISCUSSION

After the model and implementation details have been explained, different issues, regarding the efficiency, accuracy, and relevance of the models, will be addressed in this section.

Implementing efficient algorithms is crucial for atomistic simulators, which are known to stress the computer's resources. The inclusion of the charge model, and particularly the algorithm to update the local e_F and recompute the appropriate charge distribution, can significantly increase the computation time. A simple, straightforward implementation of the charge update model can increase CPU time by a factor of 10 or even more. The optimized update strategy as well as bookkeeping of the charge-modified boxes used in our implementation helped to limit the CPU time penalty just to a factor of 2.

For a relatively large simulation domain of $600\times300\times200$ nm^3 the simulator can contain as many as 10^6 boxes (2^{20}). Allocation and initialization of that many boxes along with other necessary infrastructure takes about 520 megabytes of memory. During the simulation of the subsequent process steps, even if there are high dose implants reaching amorphization level, the memory consumption does not increase much more. Simulation of the annealing, which is the most time consuming step of a kMC simulator with hundreds of billions of events, attains up to one million events per second on a 2.8-GHz CPU.

An analysis of how the simulator scales with the size of the simulation domain has been performed. The same process flow implant+annealing, including not only the charge model, but also defect diffusion, damage evolution, epitaxial recrystallization, as well as cluster and extended defect formation, was performed for different implant areas. The following process flow was used. A 10^{15}-cm^{-2} As is implanted at 50 keV. The implant temperature is 20 °C and duration is 1000 s. Then the temperature is ramped up from 20 to 1100 °C in 3 s. The implant area is ranging from 10×10 nm^2 (i.e., 10^3 arsenic ions) to 100×100 nm^2 (i.e., 10^5 arsenic ions).

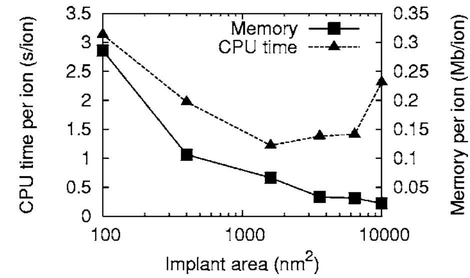

FIG. 3. CPU time and memory required per implanted ion. The same simulation, 10^{15}-cm^{-2} As implanted at 50 keV and followed by a temperature ramp from 20 to 1100 °C in 3 s, was performed for different areas of the simulation domain, ranging from 10×10 to 100×100 nm^2. Memory consumption is sublinear vs the implant area, while CPU time is superlinear for a given number of boxes. Here the number of boxes is 2^{16} for 10×10 nm^2, 2^{18} for 20×20 nm^2, and 2^{19} for all other domain sizes.

The results are shown in Fig. 3. Simulation was performed on a 2.8-GHz personal computer running Linux. The CPU time is shown in seconds per implanted As ion and memory is shown in megabytes per As ion. The nonmonotonic CPU time dependence is due to the different number of grid boxes used. For the implant area of 10×10 nm^2, there were 2^{16} boxes, for 20×20 nm^2, there were 2^{18} boxes, and for all other domain sizes, there were 2^{19} boxes.

For a fixed number of grid boxes, the CPU time increases superlinearly with the number of As particles, while the memory increases sublinearly with the number of particles. This happens because the box size increases with the size of the simulation domain and it takes more time to search for the interaction neighbors each time an event is performed. The sublinearity of the memory increase is due to a more efficient way of treating the particles based in the fact that, beyond a given concentration, it is not necessary to store the positions of all the particles.[12,13] Such a simplification does not affect the simulation accuracy, though, as the results of all simulations with the implant area of 20×20 nm^2 and larger are similar. Overall, the required CPU and memory resources look reasonable and feasible for the practical simulation of the sub-100-nm transistors on regular desktop computers.

Figure 4(a) shows the profile of substitutional boron and the hole concentration for a 900 °C simulation with an initial boron concentration of 1×10^{20} cm^{-1} from 30 to 60 nm. The boron concentration presents a profile with sharp gradients at 30 and 60 nm. The hole profile is smoother due to the Debye-length averaging explained above. The vacancy concentration profiles in their different charge states for the same simulation are shown in Fig. 4(b). The proposed charge model correctly maintains the neutral vacancies V^0 at a constant level, increasing locally the concentration of positive states V^{++} and V^+ and decreasing the V^- and V^{--} concentrations.

Figures 4(c) and 4(d) show the same boron profile simulation, but without the charge smoothing. The vacancy concentration in 4*d* is wrong due to the box discretization. In this case, even with the high boron concentration, the boxes

Downloaded 22 Sep 2005 to 157.88.111.70. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

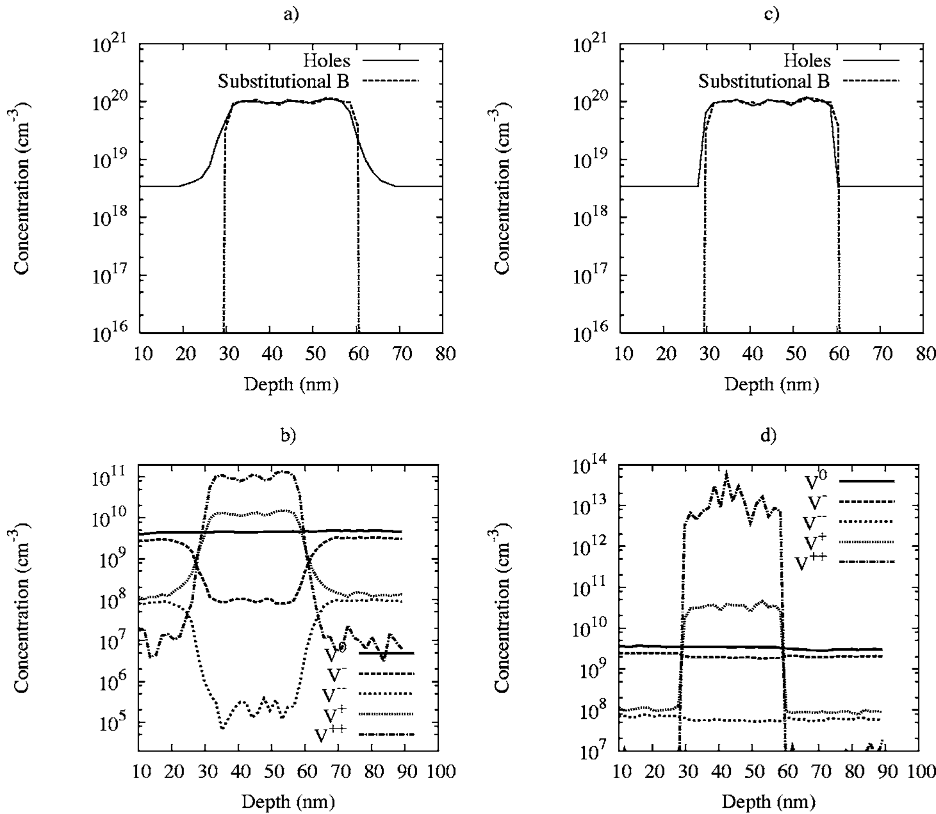

FIG. 4. (a) 1×10^{20}-cm^{-3} boron spike from 30 to 60 nm and its associated hole concentration at 900 °C. The sharp slopes of the spike are translated into the smoother slopes in the electron concentration because of the Debye-length averaging algorithm. (b) Simulated equilibrium vacancy profiles, in different charge states, for the previous simulation. Notice the constant value of V^0, a good test for the equilibrium conditions validity. The noise at low concentrations is due to the atomistic particle representation. (c) Boron spike and its associated hole concentration simulated without the Debye-length averaging algorithm. (d) Simulated equilibrium vacancy profiles, in different charge states, without averaging. This simulation clearly shows the need for the Debye-length averaging algorithm.

in the simulation are so small that they rarely contain any boron atoms, and without smoothing they report an intrinsic hole concentration. Such boxes wrongly contribute only the intrinsic concentration of V^- and V^{--} even in the high-boron-doped regions. This can be clearly seen in the Fig. 4. On the other hand, a few boxes with active dopant atoms have an unrealistically high hole concentration (that depends on the particular size of the box), which leads to artificially high V^+ and V^{++} concentrations, as can be seen in Fig. 4(d) in comparison with the profiles shown in Fig. 4(b).

Figure 5 shows two simulations of diffusion of a boron spike with a width of ≈10 nm and 2.8×10^{18}-cm^{-3} peak concentration performed at 800 °C under (a) intrinsic conditions (no extra doping) and (b) a 5×10^{19}-cm^{-3} uniform p-type background. The Annealing times (640 and 15 min) have been chosen to produce a similar diffusion length. The long hop distance (Λ),[24–26] which strongly depends on the Fermi level,[16] is ≈2.5 nm in Fig. 5(a), and ≈16 nm in Fig. 5(b). As a consequence, Fig. 5(a) follows a Gaussian distribution while Fig. 5(b) is more exponential-like, because the shape depends on Λ, as correctly predicted in the simulator model presented in this work.

Finally, Fig. 6 shows a 50-keV, 2×10^{15}-cm^{-2} As implant[27] (n-type region) over a uniformly 1×10^{18}-cm^{-3} boron-doped p-type region, after 2-h annealing at 750 °C. The experimental data (symbols) are from Ref. 28. The simulation (shown as lines) includes the amorphization and subsequent recrystallization of the amorphized silicon layer,[13] the formation of arsenic-vacancy clusters in the areas with high As concentration,[11] and the formation of extended defects. Despite the complexity of the full modeling included in Fig. 6, the apparent drift flux of B_I^0 is a direct consequence of the implemented charge model. Boron redistribution occurs during the annealing due to the charged species population difference near the pn junction and the electric-field effects. It is interesting to note that the apparent electric drift is actually associated with a net flux of neutral pairs B_I^0. The gradient of e_F produces a gradient of the relative concentration $[B_I^0]/[B^-]$ and, in turn, a diffusion flux of B_I^0 toward the n-type region, that tends to homogenize $[B_I^0]$. The boron pile up inside the n-type region is due to the low effective diffusivity of boron in the n-type regions.

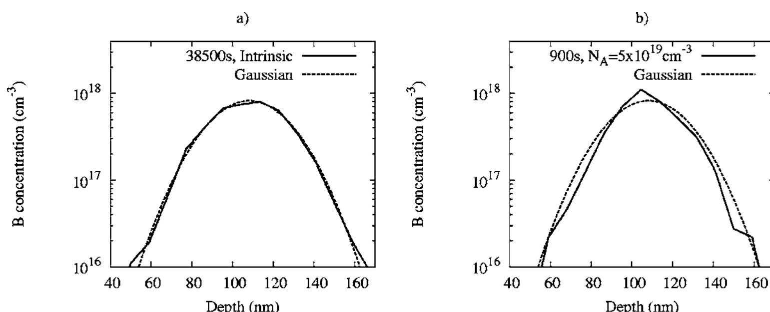

FIG. 5. Diffusion at 800 °C of a boron spike with an initial width of ≈10 nm and peak concentration of 2.8 $\times10^{18}$ cm^{-3}. (a) 640 min without background doping. The shape follows a Gaussian curve. (b) 15 min with a 5×10^{19}-p-type uniform background doping. The shape is not Gaussian. Both simulations have same diffusion length, but the shape is determined by the long hop distance, which strongly depends on e_F.

Downloaded 22 Sep 2005 to 157.88.111.70. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

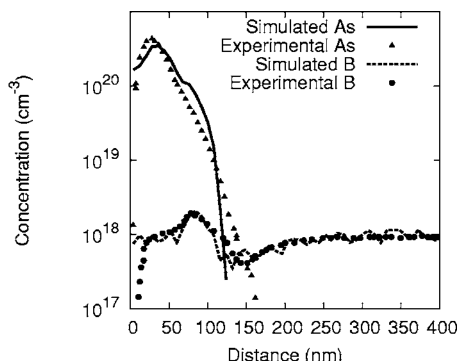

FIG. 6. Dopant redistribution in a *pn* junction. The symbols are the experimental data from Ref. 28 and the lines are our simulations. The Initially background boron concentration is 1×10^{18} cm^{-3} (*p*-type region). The n^+-type region has been implanted (As, 50 keV, 2×10^{15} cm^{-2}) and then the sample was annealed for 2 h at 750 °C.

V. CONCLUSIONS

A detailed description of the quasiatomistic model for simulating Fermi-level effects in atomistic kinetic Monte Carlo simulators has been presented. We have shown that it is necessary to smooth out the electron and hole charge distributions, and we have used the Debye length as the smoothing radius. We have discussed the major considerations for developing algorithms to update the Fermi-level-dependent parameters. The approaches used to compute the band-gap narrowing and the electric drift have been also described. The accuracy and performance of the model have been tested with simulations that illustrate the capability of the model to correctly predict the Fermi-level dependencies, and the necessity of a physically based procedure for smoothing out the carrier distribution.

ACKNOWLEDGMENTS

This work has been partially supported by the Spanish Government under Project No. BFM 2001-2250 and by the Castilla y Leon Regional Government under Project No. VA-070A05.

[1]Semiconductor Industry Association, *International Technology Roadmap for Semiconductors*, http://public.itrs.net (2003).
[2]M. Jaraiz, G. H. Gilmer, J. M. Poate, and T. D. de la Rubia, Appl. Phys. Lett. **68**, 409 (1996).
[3]M. Jaraiz, *Predictive Simulation of Semiconductor Processing* (Springer, Berlin, 2004), pp. 73–109.
[4]M. Hane, T. Ikezawa, and G. H. Gilmer, in Proceedings of the 2001 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Athens, Greece, 5–7 September 2001 (unpublished).
[5]M. Yu, R. Huang, X. Zhang, Y. Wang, and H. Oka, IEICE Trans. Electron. **E86-C**, 295 (2003).
[6]I. Martin-Bragado, M. Jaraiz, P. Castrillo, R. Pinacho, J. E. Rubio, and J. Barbolla, Appl. Phys. Lett. **84**, 4962 (2004).
[7]L. Pelaz *et al.*, Appl. Phys. Lett. **73**, 1421 (1998).
[8]L. Pelaz, G. H. Gilmer, V. C. Venezia, H.-J. Gossmann, M. Jaraiz, and J. Barbolla, Appl. Phys. Lett. **74**, 2017 (1999).
[9]L. Pelaz, V. C. Venezia, H.-J. Gossmann, A. T. Fiory, C. S. Rafferty, M. Jaraiz, and J. Barbolla, Appl. Phys. Lett. **75**, 662 (1999).
[10]W. Windl, IEICE Trans. Electron. **E86-C** 269 (2003).
[11]R. Pinacho, M. Jaraiz, J. E. Rubio, I. Martin-Bragado, and J. Barbollla, Mater. Sci. Eng., B **114–115**, 135 (2004).
[12]L. Pelaz, L. A. Marques, M. Aboy, J. Barbolla, and G. H. Gilmer, Appl. Phys. Lett. **82**, 2038 (2003).
[13]J. E. Rubio, M. Jaraiz, I. Martin-Bragado, R. Pinacho, P. Castrillo, and J. Barbolla, Mater. Sci. Eng., B **114–115**, 151 (2004).
[14]A. Asenov, M. Jaraiz, S. Roy, G. Roy, F. Adamu-Lema, A. R. Brown, V. Moroz, and R. Gafiteanu, in Proceedings of the 2002 Internatinal Conference on Simulation of Semiconductor Processes and Devices (SISPAD) Kobe, Japan, 4–6 September 2002 (unpublished).
[15]Z. Qin and S. T. Dunham, in *Si Front-End Junction Formation Technologies*, MRS Symposia Proceedings No. 717, edited by D. F. Daney, M. E. Law, A. Claverie, and M. J. Rendon (Materials Research Society, PA, 2002).
[16]I. Martin-Bragado, P. Castrillo, M. Jaraiz, R. Pinacho, J. E. Rubio, and J. Barbolla, Phys. Rev. B **72**, 035202 (2005).
[17]P. M. Fahey, P. B. Griffin, and J. D. Plummer, Rev. Mod. Phys. **61**, 289 (1989).
[18]V. Moroz, N. Strecker, and M. Jaraiz, in *Proceedings of the 32nd European Solid State Device Research Conference*, edited by G. Baccarani, E. Gnani, and M. Rudan (2002), pp. 299–302.
[19]This and other energy levels in this work are measured from the valence-band edge.
[20]C. Rafferty, in Proceedings of the 1997 International Conference on Simulation of Semicond, Processes and Devices (SISPAD), Cambridges, MA, 8–10 September 1997 (unpublished).
[21]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981).
[22]S. C. Jain and D. J. Roulston, Solid-State Electron. **34**, 453 (1991).
[23]C. Persson, U. Lindefelt, and B. E. Sernelius, J. Appl. Phys. **86**, 4419 (1999).
[24]N. E. B. Cowern, K. T. F. Janssen, G. F. A. van de Walle, and D. J. Gravesteijn, Phys. Rev. Lett. **65**, 2434 (1990).
[25]N. E. B. Cowern, G. F. A. van de Walle, D. J. Gravesteijn, and C. J. Vriezema, Phys. Rev. Lett. **67**, 212 (1991).
[26]R. Pinacho, M. Jaraiz, H. J. Gossmann, G. H. Gilmer, J. L. Benton, and P. Werner, in *Si Front-End processing: Physics and Technology of Dopant-Defect Interactions II.*, MRS Symposia Proceedings No. 610, edited by A. Agarwall, L. Pelaz, H. Vuong, P. Packan, and M. Kase (Materials Research Society, Warrendale, PA, USA, 2001), pp. B7.2.1–6.
[27]Since the as-implanted As profile (computed using an external binary collision approximation code) is shallower than the experimental profile, the subsequent annealing gives shallower As diffusion and B dip.
[28]H. S. Chao, P. B. Griffin, and J. D. Plummer, *The Influence of Amorphizing Implants on Boron Transient Enhanced Diffusion in Silicon*, MRS Symposia Proceedings No. 469 (Materials Research Society, Pittsburgh, 1997), pp. 347–352.

Downloaded 22 Sep 2005 to 157.88.111.70. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp