Atomistic Front-End Process Modelling: A Powerful Tool for Deep-Submicron Device Fabrication

M. Jaraiz*, P. Castrillo, R. Pinacho, I. Martin-Bragado, and J. Barbolla
Dept. de Electronica, Univ. de Valladolid, Spain
*Email: mjaraiz@ele.uva.es

Abstract

The complexity attained by current microelectronics process technology can hardly be handled with simulators based on the continuum approach. Over the last few years, atomistic Kinetic Monte Carlo has proven to be a new way to tackle the problems that arise as device dimension shrink into the deep submicron regime. We present some encouraging results of exploring the capabilities of this new process modelling approach.

1 Introduction

As a result of the current efforts to make progress into the deep submicron IC technology, silicon processing is facing an increasingly high level of complexity. In addition, this situation can only worsen as CMOS process technology is pushed closer to the limit of its possibilities. Front-end processing, in particular, is trying to extend the use of conventional, well established doping techniques (ion implantation plus furnace/RTP annealing) into the deep sub-micron size regime. However, many new effects show up when those techniques, well suited for larger feature sizes, are applied to sub-quarter micron processing. Since many of these effects are concurrent, the interpretation of experiments becomes ambiguous and the use of predictive process simulation becomes almost imperative. The continuum approach, based on solving partial differential equations (PDE), can barely handle the increasingly high number of different phenomena that take place, simultaneously, during deep sub-micron device fabrication. Over the last few years, we have been developing and testing the capabilities of Atomistic Front-End Process simulation based on the Kinetic Monte Carlo (KMC) technique. From our experience [1,2], Atomistic KMC process modeling seems especially suitable to fulfill these needs and is, thus, emerging as a most valuable simulation tool in the forefront of microelectronics device fabrication.

It is our purpose to show that atomistic process modeling, implemented in DADOS [3], can meet the challenge of handling, simultaneous and accurately, the wide variety of phenomena involved in current deep submicron front-end processing.

2 Setting the Problem: What do we want to model?

Figure 1 represents the formation of the source/drain (S/D) region of a 0.1um NMOSFET, after a short anneal. An amorphous region created by the As implant has recrystallize very quickly, leaving behind only the foreign impurity atoms.

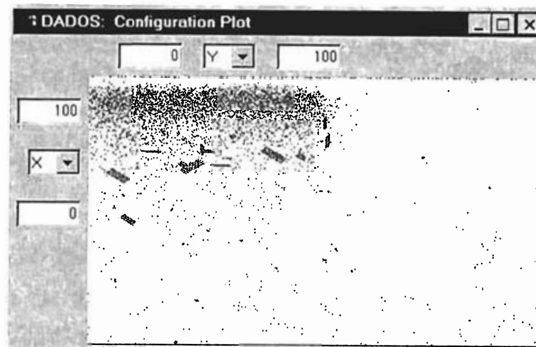

Figure 1. Typical view of a S/D extension implant. {311} defects can be seen beyond the a-c interface after recrystalization of the amorphised region.

Excess I beyond the a-c interface have formed small clusters and some {311} defects (and in some cases dislocation loops) that are now dissolving and producing TED, as well as forming I-B complexes with the B atoms present in the channel. The very high As concentration induces Fermi level effects on diffusion. Moreover, due to the dimensions involved, 2D or 3D simulation is needed to account for realistic effects. Finally, the discrete nature of the dopants and point defects together with their possible charge state can introduce particular correlations in their spatial distribution.

In what follows, we will briefly show how KMC Atomistic Process Simulation can implement these phenomena, in a way accurate and efficient enough to tackle the challenge of including *simultaneously*, in a *3D simulation* all of the mechanisms and highly *non-equilibrium* processes mentioned above. Since we have already presented some of the features of DADOS, we will describe here mainly mechanisms recently incorporated, namely Fermi level effects, and amorphization-recrystallization.

3 The KMC Concept

KMC, unlike Molecular Dynamics, is an event-driven technique, i.e., simulates events (e.g. diffusion hops) at random, with probabilities according to their respective event rates. In this way, it self-adjusts the timestep as the simulation proceeds, just to be able to account for the fastest event present at the current time.

Figure 2 illustrates the concept behind the KMC approach: Fig. 2a (from Ref. 4) shows a high resolution TEM view of a silicon sample with a {311} extended defect embedded in the silicon atom rows. In MD, all lattice and defect atoms are simulated. In KMC, only the atoms belonging to point or extended defects (represented as circles on the TEM view in Fig. 2b) are simulated. In the actual sample, all lattice atoms would be vibrating (with a period of about 10^{-13} s) and, from time to time (e.g. every 10^{-9} s), one of the

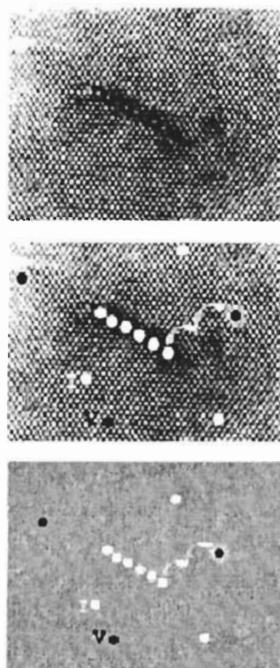

Figure 2. Sketch of the KMC concept.

isolated point defects would jump to a neighboring position, where it might be captured by an adjacent extended defect. At even longer time intervals, (e.g. every 10^{-3} s) a point defect would be emitted from the extended defect. To simulate this, KMC (Fig. 2c) would start out with timestep of about 10^{-9} s, since it only follows the defect atoms. In addition, the fast moving point defects disappear very quickly, leaving only the extended defect and allowing the timestep to automatically be raised to 10^{-3} s.

3.1 KMC Basic Components

A KMC simulator consists of a 3D Simulation Box, of dimensions ranging from tens of nanometers to a few microns, where point and extended defects can be created and allowed to perform events (point defect jump, emission from a cluster,) and interact with each other (annihilation, capture by a cluster,). Point defects can be created in the simulation box to simulate, for example, an ion implantation step. For this purpose, a BCA (Binary Collision Approximation) code can be used to calculate the coordinates of the I and V generated by each ion (cascade). Alternatively, V and I can be generated and recombined at one of the box surfaces to simulate a free surface

DADOS is a non-lattice KMC simulator. The main reason for this is that the processes and dimensions we want to be able to simulate do not need (i.e., are insensitive to using) a lattice. Besides, the computer requirements would be prohibitive for many practical situations. Lattice KMC is necessary, for example, to simulate grain boundaries in polycrystalline materials [5].

4 Atomistic KMC modeling of Diffusion and Defects in Si: a Brief Review

Ion implantation is routinely used in current silicon technology for different purposes like doping and amorphization. Ion implantation damages the crystal lattice. From a practical standpoint for microelectronics processing, implantation damage can be classified into two categories depending on whether or not amorphization is reached. Here we focus on the situation where the damage is not enough to amorphize the lattice. In that case, subsequent high temperature treatments transform the lattice damage into separate, well localized extended defects by either the agglomeration of point defects or by the anneal of high disorder regions.

The basic defects in Si are the intrinsic point defects: self-interstitials (I) and vacancies (V). In addition, there is always a certain amount of unwanted impurities, particularly carbon and oxygen, which form the native extrinsic defects. Finally, other impurities (boron, phosphorous, arsenic) are intentionally introduced to dope or otherwise modify the Si properties. Their presence adds new defect forms and interaction mechanisms. A generic picture of the basic interaction of a

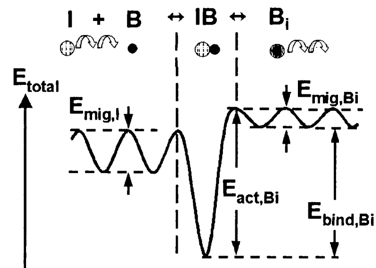

Figure 3. Generic interaction of a Si point defect (I) with an impurity atom (B). Energies involved in the process can be calculated by ab-initio methods.

point defect (I, V) and a foreign atom is shown in Figure 3. In this example, a migrating free I encounters a substitutional impurity atom B and binds to it forming an IB pair that may be mobile or not. The energies involved (migration, binding...) can often be calculated by ab-initio methods.

The agglomeration of V and I can lead to the formation of Si extended defects (like I {311} and dislocation loops, and V voids). Small I clusters are likely to have a non-monotonic behavior in terms of both shape and energy, as inferred from calculations and from inverse-model analysis of diffusion data (Figure 4a, from Ref. 6). Energy oscillations are important because they can act as a bottleneck during the Ostwald ripening process by which small, unstable clusters dissolve while the big, more stable ones grow. The {311} 'rod like' defect has been a subject of intense research over the last few years [7] because of their close association with transient enhanced diffusion (TED)[4]. If the dose and temperature are high enough another form of extended defect, the dislocation loop (DL), appears. Knowledge of geometry and energies (Figure 4b, from Ref. 8) permits direct implementation of these defects into an atomistic process simulator, as it has been recently done in DADOS.

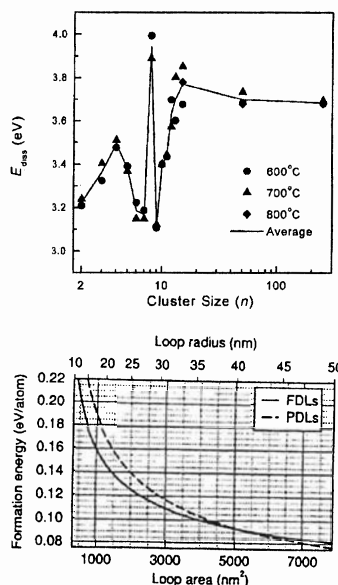

Figure 4. (a) Energies of {311} defects (from Cowern et al. Ref. 6) and (b) Energies of dislocation loops (from Cristiano et al. Ref. 8)

Overall, there is a continuous defect morphology evolution throughout a high temperature annealing stage. Mobile point defects very soon agglomerate forming small clusters. Then the Ostwald ripening process begins: less stable clusters (higher emission/capture ratio) dissolve while the more stable ones grow and evolve into {311}. If the free surface is far away and the dose is high enough the {311} can evolve into the more stable faulted DL and these, finally, into the perfect DL which have a lower formation energy per atom than the faulted DL beyond a threshold size.

Small vacancy defects have been experimentally identified by electron paramagnetic resonance. Vacancy clusters large enough to be seen by electron microscopy appear as spherical voids. Similar to the I clusters behavior just described, when a small V cluster grows beyond a given size it is re-shaped into a void defect. This is necessary to maintain the correct volume to surface ratio as the V cluster grows. A large cluster of N vacancies is re-shaped to be spherical, occupying the volume corresponding to the same number of Si lattice sites.

Under certain conditions, the mechanisms given so far are not enough to explain some experimental observations. For example, it is known that, above a certain concentration, dopants become electrically inactive. It has also been shown that a high I concentration can render electrically inactive a fraction of substitutional B even at a concentration below its solid solubility [9]. An I_nB_m clustering mechanism has been invoked to account for these facts [2]. These extended defects, can be modeled in a way parallel to the clusters, but with two lists of constituent

atoms. An analogous modeling has also been successfully employed for carbon, as shown in Fig. 5 [10].

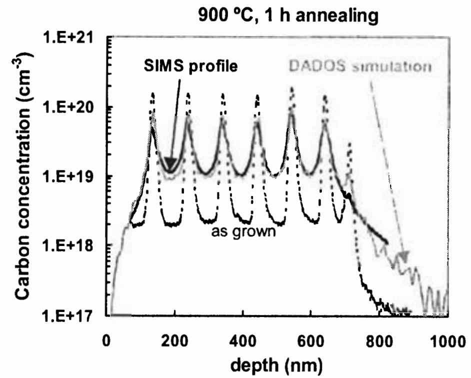

Figure 5. Experimental SIMS and DADOS-simulated carbon profiles in silicon after 1 hour annealing at 900C.

Electrical deactivation can also be due to precipitation. Small X-I clusters might be the embryos of SiX precipitates and clustering and precipitation occur together with one process dominating over the other depending on the experimental conditions.

The models presented above can also account for some forms of gettering. Solid solubility effects have also been implemented in DADOS for Carbon, as an example. When C particles are captured by (or 'deposited at') the surface 'extended defect', it stores them and emits C_i point defects at a rate proportional to the C areal density up to a limit that is related to the solid solubility of C in Si.

5 Fermi Level Effects and Amorphization

Fermi level effects are implemented in DADOS as follows. The ratio between different charge states for a given species can be assumed to respond instantaneously to the local electron concentration. Every certain number of events or elapsed time, the atomistic configuration is analyzed to generate a set of data that includes the current 1D, 2D or 3D concentration profiles, cluster size histograms and so on. We shall refer to this set of data as a time frame. The time interval between time frames is chosen such that they change smoothly from one to the next. The electron concentration profile (1D, 2D or 3D) is updated every time frame and used to generate a corresponding transition probability profile for each species to each charge state. Once the time-frame data set is built, the simulation proceeds and, every time a point defect jumps, the charge state is updated. In addition, the effect of the electric field on charged point defects has to be taken into account. The presence of a macroscopic electric field (weaker than the local electric field) is assumed not to change the jump distance or jump rate, but to simply introduce a bias in the diffusion jump direction probabilities. Figure 6 shows a test example of charged V profiles.

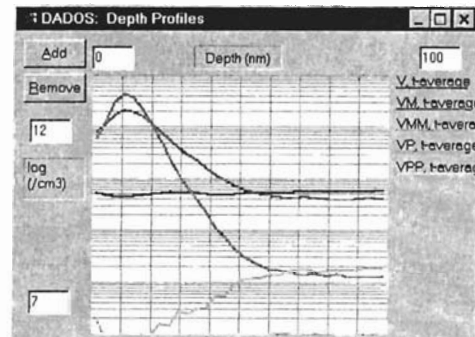

Figure 6. Depth profiles for different V charge states in an As doped region. The flat profile of neutral V indicates that the balance between electric field and charge population is correctly achieved by the atomistic model described in the text.

The global nature and behavior of implantation damage in silicon is a complex issue and is currently the subject of active research. For instance, as the temperature is raised clusters are evidenced experimentally [11] as V-voids and I-dislocations (Fig. 7a). This damage pattern (V clusters near the surface, I clusters deeper) can be readily understood taking into account the ion momentum. However, chemical effects must also have played a critical role, as it is obvious from the comparison of Fig. 7a and Fig. 7b, which include pairs of ions with almost the same mass but one is a dopant (As) while the other is isoelectric (Ge). Similar results have been obtained for P and Si [11].

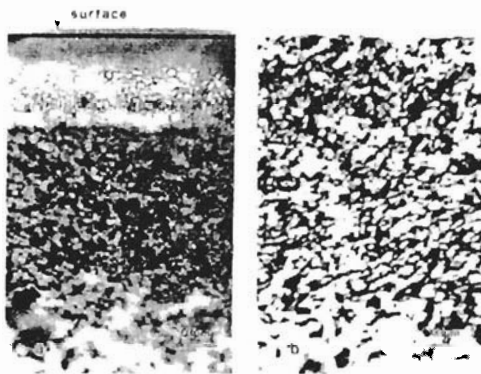

Figure 7. (From Holland et al. Ref. 11) Chemical effects have to be invoked to explain the different damage morphology of two atoms of almost the same ion mass, As (left) and Ge (right). In Atomistic modeling it is easy to specify interaction rules depending on the type of atom.

Amorphization, unlike the extended defects considered up to here, is a massive form of lattice disorder. As such, it is not amenable to the atomistic modeling employed so far. Yet, it has been recently incorporated within a KMC scheme in DADOS as follows.

For simplicity we shall refer to only one dimension histograms, but it is implemented in 3D with 2×2×2 nm^3 boxes. Figure 8 shows a damage depth profile at two consecutive time frames t1, t2, during an implant. The damage depth profiles correspond to the I and V profiles generated by the implanted ions. Normally, the I and the V profiles are indistinguishable from each other because their values are much larger than their difference. Indicated in the figure are the damage threshold for amorphization and a 'MaxStorage' level. The level for amorphization is typically between 20%-30% of the Si atomic density. The MaxStorage level is defined as the maximum allowable I and V concentration in the simulation. It is chosen much lower than the amorphization level, to keep the number of particles in the simulation manageable, but high enough to still reproduce dynamic annealing effects reasonably well. Typically, the MaxStorage level is two orders of magnitude lower than the amorphization level.

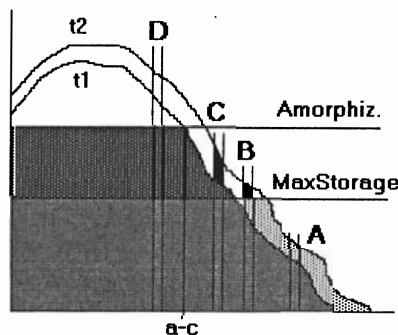

Figure 8. Schematics of the dynamic annealing, amorphization and recrystalization as implemented in the KMC simulator DADOS (see text).

At depth A, the damage level is low and all I and V particles generated by the cascades are incorporated into the atomistic simulation. At depth B and time t2, the damage has reached past the MaxStorage threshold: the number of I and V particles in excess are not inserted into the simulation box, but the damage histogram is updated. Likewise, at depth C, time t2, all I and V particles above the MaxStorage level are deleted, and the damage profile is updated. Finally, at depth D (already amorphised), a high concentration of I and V particles equal to MaxStorage is kept, but the region is labeled as 'amorphous'. Notice that, since there is dynamic anneal between t1 and t2, it is possible for regions B or C to go back to a damage level below the MaxStorage threshold. This, however, is not allowed to occur in region D that can only re-crystallize from the amorphous-crystalline (a/c) interface. The process of deleting I and V particles in excess of the MaxStorage limit always takes care of a particular requirement: the local imbalance between the I and V profiles, due to the ion momentum (see Fig. 8), has to be maintained. V and I particles generated by the cascades are inserted into the simulation as point defects. However, upon interaction, I and V point defects do not directly recombine but form an IV pair point defect with a barrier for recombination. In addition, I and V are assumed to initially form 'amorphous' agglomerates, similar to the clusters described above, but with a lower activation, in agreement with experimental observations [12,13]. Recrystalization is implemented as a special event type. The a/c interface advances into the amorphous regions at a temperature dependent velocity. The recrystalization procedure deletes all I and V particles, only impurity atoms are left behind.

The accuracy of the above described method is quite remarkable. An example can be seen in Fig. 9, in which DADOS correctly predicted that a 50 keV, 3.6 cm^{-2} Si implant will amorphize up to slightly beyond 100 nm, but not the 5 nm adjacent to the surface, in agreement with the experiment [14].

All of the mechanisms described above that are implemented in DADOS are active in all simulations without degrading computation speed noticeably. For example, on a 400 MHz Pentium-II it simulates more than 600 000 events per second, including charge update and for any number of different impurity atoms.

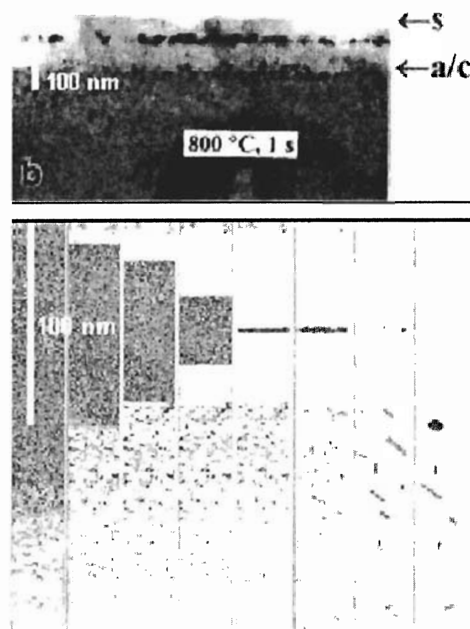

Figure 9. Top: (from Pan et al. Ref. 14). XTEM view showing recrystallized layer and defects at the center and beyond the a/c interface. Bottom: Sequence, from a DADOS simulation, predicting the correct amorphization and annealing behavior.

6 Conclusions

In addition to the accurate modeling of diffusion and defects, already demonstrated in atomistic KMC, we have recently included Fermi level effects and amorphization. This brings the KMC approach to a fully operational front-end process simulator stage, particularly efficient and realistic under extreme conditions such as deep submicron dimensions, complex processing scenarios and highly non-equilibrium situations. In summary, atomistic front-end process modeling has attained a level of maturity that enables it to advantageously simulate the processing steps of current deep submicron device technology.

Stimulating discussions with Drs. C. S. Rafferty, G. H. Gilmer and M. Hane are appreciated.

7 References

[1] M. Jaraiz, G. H. Gilmer, J. M. Poate, and T. D. de la Rubia, Appl. Phys. Lett. 68 (1997) p. 409.
[2] L. Pelaz, M. Jaraiz, G. H. Gilmer, H. J. Gossmann, C. S. Rafferty, D. J. Eaglesham, and J. M. Poate, Appl. Phys. Lett. 70 (1997) p. 2285.
[3] DADOS (Diffusion and Defects, Object-oriented Simulator), M. Jaraiz, L. Pelaz, E. Rubio, J. Barbolla, G. H. Gilmer, D. J. Eaglesham, H. J. Gossmann, and J. M. Poate, Mat. Res. Soc. Symp. Proc. 532 (1998) p.43.
[4] D. J. Eaglesham, P. A. Stolk, H. J. Gossmann, and J. M. Poate, Appl. Phys. Lett. 65 (1994) p. 2305.
[5] J. E. Rubio, M. Jaraiz, L. A. Bailon, J. Barbolla, M. J. Lopez, and G. H. Gilmer, Mat. Res. Soc. Symp. Proc. 514 (1998) p.127.
[6] N.E.B. Cowern, G. Mannino, P. A. Stolk, F. Roozeboom. H. G. A. Huizing, J. G. M. van Berkum, F. Cristiano, A. Claverie, and M. Jaraiz, Phys. Rev. Lett. 82, (1999) 4460.
[7] S. Takeda, Jap. J. Appl. Phys. 30 (1991) p. L639.
[8] F. Cristiano, J. Grisolia, B. Colombeau, M. Omri, B. de Mauduit, A. Claverie, L. F. Giles, and N. E. B. Cowern, J. Appl. Phys., 87 (2000) p. 8420.
[9] P. A. Stolk, H. J. Gossmann, D. J. Eaglesham, D.C. Jacobson, J.M. Poate, and H.S. Luftman, Appl. Phys. Lett. 66 (1995) p. 568.
[10] R. Pinacho, M. Jaraíz, H. J. Gossmann, G. H. Gilmer, and J. L. Benton, Mat. Res. Soc. Spring Meeting 2000. (In press).
[11] O. W. Holland, L. Xie, B. Nielsen, and D. S. Zhou, J. Electronics Mat., 25 (1996) p. 99.
[12] P. J. Schultz, C. Jagadish, M. C. Ridgway, R. G. Elliman, and J. S. Williams, Phys. Rev. B, 44 (1991) p.9118.
[13] R. G. Elliman, J. Linnros, and W. L. Brown, Mat. Res. Soc. Symp. Proc., 100 (1988) p. 363.
[14] G. Z. Pan, K. N. Tu, and S. Prussin, Appl. Phys. Lett., 71 (1997) p.659.