The Process Modeling Hierarchy: Connecting Atomistic Calculations to Nanoscale Behavior

Scott T. DUNHAM*, Pavel FASTENKO†, Zudian QIN§, and Milan DIELBEL‡,

SUMMARY In this work, we review our recent efforts to make effective use of atomistic calculations for the advancement of VLSI process simulation. We focus on three example applications: the behavior of implanted fluorine, arsenic diffusion and activation, and the impact of charge interactions on doping fluctuations.

key words: atomistic, diffusion, activation, doping fluctuations, silicon, arsenic, fluorine

1. Introduction

Accurately predicting nanoscale device structures resulting from advanced integrated circuit fabrication processes is a very challenging problem due to the large number of strongly interacting reactions, the dominance of non-equilibrium behavior, and the fine tolerances required. Great advances have been made in understanding ion implantation, defect-mediated dopant diffusion, extended defect kinetics and film growth along with their interactions, but many mysteries still remain. Atomistic techniques provide a range of powerful tools including ab-initio calculations based on density functional theory (DFT), empirical-potential molecular dynamics (MD), and kinetic Monte Carlo (MC). These tools have the potential to significantly advance the predictive capability of process simulators. However, substantial challenges must be overcome to make effective use of such calculations due to the wide gap which is present in the spatial and temporal scales.

The accuracy of these calculations is still generally insufficient for their direct use in quantitative prediction of device structures. However, as we demonstrate with the example of implanted fluorine in silicon, DFT calculations are very effective in identifying the most likely reaction paths and dominant or rate-limiting processes among the complex array of possibilities. We then use this information to develop continuum models which are able to predict unusual features of experimental behavior.

Next, we discuss the diffusion and activation of arsenic. This analysis utilizes both direct DFT calculations as well as kinetic lattice Monte Carlo (KLMC) simulations to gain understanding about arsenic behavior, which is then implemented in continuum models. Because it considers systems at the atomic level, but can be scaled up to consider nanoscale behavior at long time scales (seconds), KLMC is a particularly effective tool when complex cooperative effects or local correlations are important.

Finally, we consider direct application of atomistic simulation via KLMC. As an example, we investigate doping fluctuations which lead to threshold voltage variations by including Coulombic interactions and charge redistribution in KLMC simulations.

2. Behavior of Implanted Fluorine

Fluorine has been shown to have beneficial properties on both boron TED reduction [1]-[4] as well as boron activation [1], [2]. However, to utilize these benefits effectively, a fundamental understanding of F behavior is necessary, particularly since implanted F shows anomalous diffusion behavior [5]: instead of broadening, annealed profiles first remain immobile and then sharpen and shift toward the surface (Fig. 2).

Based on DFT calculations [6], [7], a single F atom prefers to reside in a bond centered interstitial site. This site is preferred by 0.18 eV over the tetrahedral interstitial configuration and by 1.0 eV over the lowest substitutional site. Interstitial F (F\textsubscript{i}) is highly mobile (E\textsubscript{m} = 0.73 eV), so rapid diffusion would be expected even at low temperatures, in contrast to experiments (see Fig. 2). Thus, F must form more stable clusters.

Via exploration of potential structures via DFT, no significant binding of F\textsubscript{i} to each other or to self-interstitials was found. However, we found F\textsubscript{n}V\textsubscript{m} struc-
Fig. 2 Comparison of simulation (bottom) with experimental data [5] (top) for $10^{18}$cm$^{-2}$ 30 keV F$^+$ implant annealed for 30 min at various temperatures. The simulation includes $F_nV$, $F_nV_2$, and FI clusters in addition to an interstitial clustering model [8]. The initial defect/fluorine profiles were obtained with MC ion implant simulation (UT-Mariano).

Fig. 3 Binding energy of $F_nV_m$ configurations relative to interstitial fluorine ($F_i$) and single vacancies ($V$). The decreasing binding energy of the $F_nV$ structures is attributed to the increasing crowding of the F atoms.

Fig. 4 Equilibrium concentration of various $F_nV_m$ structures vs. total F concentration at 600°C. For low $C_{F_m}$ the dominant species is $F_i$ due to the entropy of mixing. At high $C_{F_m}$, F in Si is primarily in the form of $F_nV_2$ clusters. Note that the vacancy formation energy is included in these calculations.

structures to have high binding energies (see Fig. 3) [6], [7]. Note that this notation considers $F_i$ as the reference structure since it is the most stable. The structure of nominally substitutional F supports this choice as rather than sitting symmetrically, F moves close (<1.7 Å) to one Si neighbor, as though passivating a dangling bond associated with vacancy. A vacancy or V cluster can potentially bind as many F atoms as the number of dangling bonds (e.g., 4 F for V, 6 F for $V_2$).

For two or more F atoms, $F_nV_m$ clusters are energetically favorable even when the formation energy of V (3.4 eV) is considered. The saturated $F_6V_2$ structure is even stable in the presence of interstitials (binding energy exceeds Frenkel pair formation energy). Because of this strong binding, F can be expected to reside in $F_nV_m$ structures rather than as isolated $F_i$ as shown in Fig 4. $F_nV_m$ will dominate much more strongly during initial stages of implant anneals due to the presence of excess vacancies.

When $F_i$ diffusion and fluorine decoration of V and V$_2$ is included in a continuum model along with point defect diffusion and clustering/extended defect models [8] and F segregation to and diffusion through the thin oxide film, the resulting simulation profiles correctly predict the apparent uphill diffusion observed experimentally during annealing of implanted F (Fig. 2(b)) [6], [7]. This behavior is due to the $F_nV_m$ dissolving in the deeper I-rich regions, but remaining in the shallow V-rich regions, as fast-diffusing $F_i$ are trapped by $V_m$ and released by I.

Our DFT results find no significant B/F clusters [6], [7]. However, they do suggest a possible source for the effect of F on B shallow junction formation via modification of point defect concentrations. For non-amorphizing F implants, the implanted F can be considered to result in interstitial F (any substitutional F will be matched with an equal number of excess I). However, unlike substitutional dopants, the F does not want to become substitutional (which would give +1 excess self-interstitials). Instead, it wants to form $F_nV_m$ clusters. Formation of the dominant $F_6V_2$ cluster (and
also the $F_3V$ cluster) leaves one excess I for each three F (+1/3), resulting in a small increase in TED (the effect is actually less than +1/3 due to F diffusion removing interstitials). In contrast, F implanted into amorphous material results in reduced TED. The strong $F/V$ affinity is expected to lead to V incorporation as F passivates bonds in the amorphous material and then incorporates as $FV$ (substitutional F) or $F_nV$ during regrowth. As the F then forms larger, more stable clusters and diffuses to surface, V are released (e.g., $6FV \Rightarrow F_3V_2 + 4V$). The consequence for boron diffusion is that these excess V reduce the number of I and thus reduce TED and enhance activation for pre-amorphized samples, consistent with experimental results [1]–[4].

3. Arsenic Diffusion and Activation

Formation of ultrashallow As junctions is critical to the future of CMOS technology. Understanding As behavior is a substantial challenge because of strong interstitial injection due to deactivation [9],[10], rapid deactivation at low temperatures [11], and greatly enhanced diffusion at high doping levels [12]. Our modeling strategy is to use atomistic methods to develop understanding that can then be translated into fundamental models at the continuum level.

There is a strong attractive interaction between arsenic and vacancies due to the combination of strain compensation and valence. As a result, under most conditions deactivation of arsenic occurs primarily via the formation of arsenic vacancy complexes. DFT calculations [13],[14] find that $As_3V$ complexes (4 substitutional As atoms surrounding an empty lattice site) are the most energetically favorable, which seems reasonable given that this configuration allows each valence 5 As to have 3 nearest neighbors. The activation of As as a function of total concentration is consistent with $As_3V$ as dominant cluster (Fig. 5) [16].

It has been observed experimentally [11],[19] that for high active arsenic concentrations, initial deactivation is very rapid (within 15 seconds) even at low T (down to 500°C, Fig. 7). The deactivation of high concentration arsenic layers also injecting interstitials into the substrate [10] (Fig. 8). The problem associated with this data is where the vacancies come from. Diffusion from the surface or normal bulk Frenkel pair (I+V) generation are both much too slow. A possible answer (previously suggested by Parisini et al. [15]) is that Si atoms are ejected from $As_nSi$ clusters, forming an arsenic-vacancy cluster and a self-interstitial. Ab-initio calculations suggest [20] that energetically the most favorable reaction is: $As_3Si \rightarrow As_2V + I$.

We have tested this proposition by calculating the energy barriers for ejection of a silicon atom from an $As_3Si$ tetrahedral cluster via DFT and the nudged elastic band (NEB) [21] method. The results of the NEB calculation are shown in Fig. 6 [16]. The overall process is particularly interesting because the presence of the arsenic atoms allow for the creation of an interstitial vacancy pair with a barrier of 1.4 eV which is much less than the DFT Frenkel pair formation energy of more than 7 eV.

![Fig. 5](image5.png) Comparison of arsenic-vacancy clustering model to experimental measurements of active versus total arsenic concentration at long times. The model was fitted to data from Solmi et al. [17] at 1100°C and Guerrero et al. [18] at 950°C and extrapolated to lower temperature. Extrapolation to 750°C temperature is compared to data from Rousseau et al. [9].

![Fig. 6](image6.png) Energy versus distance for kick-out of a silicon atom from between 4 substitutional As atoms.

Since there is a relatively low barrier for I ejection, in heavily As doped material, the clusters can generate their own vacancies, resulting in initial rapid deactivation. This process is highly nonlinear since the density of $As_3Si$ in equilibrium is proportional to $C_{As^4}$. The formation of new $As_3Si$ structures requires random As migration. Since V are trapped by As clusters, I-mediated diffusion dominates. Thus, the process is self-accelerated by the I injection: cluster formation...
generates I which enhance As diffusion and precursor cluster formation, which in turn enhances I injection. We incorporate these concepts into a continuum model [16, 22, 23] and the result (as shown in Fig. 7) is rapid deactivation even at very low T (500°C). This system also predicts the I injection kinetics. We used a moment-based {311} loop model [24] to account for extended defect formation. Figure 8 shows comparison of simulation results to experimental data of Rousseau [9], [10] for arsenic layer which was initially fully-activated via laser anneal.

![Graph showing deactivation kinetics for laser annealed As layers.](image1)

**Fig. 7** Deactivation kinetics for laser annealed As layers, with comparison of simulations to experimental data from Luning [11]. Rapid deactivation (<15 s) is observed even for the lowest T (500°C).

Experiments find that arsenic diffusion is dramatically enhanced at doping concentrations above 10^{19} cm^{-3} [12]. At such high concentrations, complex interactions between multiple arsenic atoms and point defects can be expected. Kinetic lattice Monte Carlo (KLMC) provides an effective tool to investigate the behavior of such systems as the method retains the underlying atomic structure and mechanisms and yet enables consideration of system sizes and process times appropriate to submicron device fabrication [25]–[27]. DFT calculations indicate that there is a long-range dopant/vacancy attraction [28], [29], leading to interaction of V with multiple As atoms at high doping levels. Using DFT interaction energies in KLMC simulations leads to strong enhancement of D_{As} at high concentrations [25]–[27], consistent with experiment (Fig. 9).

![Graph showing comparison of KLMC predictions to experimental data.](image2)

**Fig. 9** Comparison of KLMC predictions [25]–[27] to experimental measurements of arsenic diffusivity at 1050°C versus doping level [12], [30]. The diffusivity for moderate doping levels is fit assuming diffusion via negatively-charged vacancies and then extrapolated to higher doping levels using the KLMC simulation results. Note a sharp increase in arsenic diffusivity for donor concentrations exceeding 2 \times 10^{20} cm^{-3}.

A key feature of As diffusion profiles is a mismatch between active doping level and the much higher (by order of magnitude) apparent solubility associated with shoulder of diffusion profile [31] (Fig. 10). Although DFT calculations indicate that As deactivation is due to immobile As_{n}V clusters [13], [14], [20], the high concentration depletion effect illustrated in Fig. 9 can account for this otherwise anomalous behavior. Because of the dramatic increase of diffusivity at high doping levels, the small gradient in free As concentration associated with As_{n}V clusters (Fig. 5) leads to a substantial diffusion flux.

![Graph showing arsenic and boron concentration profiles.](image3)

**Fig. 8** Experiment by Rousseau et al. [10] investigated As deactivation and associated I injection. A laser anneal was used to initially fully activate heavily-doped As region, and underlying boron buried layer monitored I supersaturation during subsequent annealing. Comparison is shown for 2h anneal at 750°C to model predictions which capture both deactivation kinetics and enhanced B diffusion.

A simple model also works successfully for shallow junctions once the initial transient is over, as illustrated in Fig. 11. However, when we use this model to predict the initial diffusion behavior following implantation, we greatly overestimate diffusion (Fig. 12(a)) [22], [23]. The large junction motion is caused by interstitial...
injection due to As$_4$V formation. We can understand this behavior by noting the strong As/V binding. During regrowth, this can be expected to stabilize V incorporation in regrown region [22],[32]. The effect is similar to what we propose for F, but we assume at least 2 neighboring As atoms are required since As/V binding is weaker than F/V binding. The grown-in V allows rapid As deactivation without too much diffusion due to I injection from As$_3$Si clusters [22],[23], leading to accurate prediction of junction motion (Fig. 12(b)). This picture is supported by positron annihilation experiments [34] which find a large vacancy population in laser annealed samples, and by reduction of {311} defects and loops in high dose low energy As implanted silicon [35].

4. Charge Interactions and Doping Fluctuations

Reduced device dimensions and thermal budgets greatly reduce computation times for KLMC simulations, while requiring the inclusion of complex nonequilibrium behavior that makes continuum simulations more computationally expensive. Thus, KLMC simulations become a realistic alternative for simulation of deep submicron structures. We have previously demonstrated 3D KLMC simulations of sub-100nm nMOS structures [26]. Such atomistic simulations become necessary for deep submicron devices because the modest number of dopants and defects present lead to significant variations in device properties. To accurately model these effects, it is necessary to include electric field effects on dopant redistribution, which have previously been neglected in atomistic simulations [36],[37].

Charge effects are modeled in continuous simulations through the quantity of local carrier concentration, which is normally calculated from dopant profiles using the charge neutrality assumption. Unfortunately, dopant concentration is no longer a valid concept at the atomic scale, where dopant atoms are considered to be discrete and point-like. Each donor (acceptor) contributes an electron (hole) cloud in its neighborhood. The classical solution for carrier density near a charged dopant in a semiconductor (screened Coulomb potential) diverges as distance r approaches zero, which is physically incorrect and causes a problem in the carrier concentration calculation. To overcome this problem, we use a quantum calculation of charge density (quantum perturbation method). We find that the resulting distribution matches the classical screened Coulombic solution at large distances, but is very different at short range (<1nm) as it avoids the singularity at r = 0 [38],[39] (Fig. 13). Upon examination of numerical results for different background carrier concentrations and temperatures, we find that the quantum solution can be modeled by (Fig. 14):

$$\rho(r) = \rho(0) \frac{r_0}{\sqrt{r^2 + r_0^2}} \exp \left( -\frac{r}{L_D} \right)$$

where $\rho(0)$ is the charge density at the origin, $L_D$ is the Debye length, and $r_0 \approx 0.8$ nm is a weak function of the background carrier concentration and temperature. Note that for $r_0 \rightarrow 0$, this expression reduces to the classical solution.

Given the charge distribution in the neighborhood
within KLMC at an elevated temperature and the electron potential distribution again calculated at 300K.

Fig. 16 compares resulting histograms, which show that the system has a narrower potential distribution after annealing [38],[39]. We attribute this to dopant/dopant repulsion leading to a more uniform dopant distribution (no narrowing in potential distribution is seen if effect of potential on charged defects is neglected). This is favorable for further device scaling. Vth is limited by the lowest energy path between source and drain in plane of inversion layer. As seen in Fig. 16, the low energy tail of the potential distribution is strongly reduced due to ordering, so smaller Vth variations will result.

Fig. 12 A misc SIMS profile for 5keV 2 x 10^15 cm^-2 implant followed by spike anneal at 950°C. (100°C/s ramp-up rate). Data from Jain [33]. Also shown are model predictions. In (a), the regrown region is assumed to be defect-free, while in (b) the vacancies are assumed to be incorporated on sites which have two or more nearest neighbor As atoms.

of a single point charge, the local carrier concentration is determined by summing contributions from all ionized dopant atoms and charged defects. The resulting electrical potential then enters into the hopping rate calculations for all charged defects, giving appropriate drift terms [25]–[27]. One example used to test this implementation is shown in Fig. 15. Multiple dopants diffusing simultaneously interact via field effects. Initially, there is a uniform boron background and non-uniform As doping in the system. After the simulation, B is swept out of the junction region where the field gradient is the strongest.

Channel dopant fluctuations are a concern for nanoscale devices as they impact transistor characteristics such as threshold voltage (Vth). To analyze the effect of charge interactions on doping fluctuations, atoms are first randomly initialized in the system, and the carrier distribution and associated electrical potential calculated for T = 300K. The system is then annealed

5. Conclusions

We have illustrated some of the ways in which atomistic calculations can be effectively used to advance VLSI process modeling. There is clearly great potential for such methods to play an increasingly important role in advancing silicon technology.

6. Acknowledgements

This paper describes the efforts of many collaborators, including Hannes Jónsson, Blas Uberuaga (now at LANL) and Graeme Henkelman from the University of Washington Department of Chemistry, and Amitabh Jain and Sriniv Chakravarthi from Texas Instruments. The work was supported primarily by the Semiconductor Research Corporation. Many of the calculations were conducted using a computing cluster donated by Intel.
Fig. 14 Excess electron distribution in the neighborhood of a positive point charge for system with background electron concentration of $1 \times 10^{19}$ cm$^{-3}$ (T=1000$^\circ$C). The quantum solution matches the classical solution at long range. Also plotted is Eq. (1), which shows excellent agreement with the quantum solution over the entire range.

Fig. 15 KLMC simulation shows coupled diffusion due to field effects. B is swept out of the junction region by the presence of strong electric field.

References

[33] A. Jain, SIMS and resistivity data from Texas Instruments.

**Scott Dunham** was born in Brattleboro, Vermont in 1968. He received his B.S. degree from Cornell University (Ithaca, NY) in 1989 and his M.S. and Ph.D. degrees from Stanford University (Stanford, CA) in 1989 and 1990, all in electrical engineering. He was a faculty member in the Electrical and Computer Engineering Department at Boston University from 1985 to 1999. He is currently Professor of Electrical Engineering and Adjunct Professor of Material Science and Engineering and Physics at the University of Washington (Seattle, WA). The focus of his research is on process and device modeling.

**Pavel Fastenko** was born in Moscow in 1973. He earned his Diploma in electrical engineering in 1996 from Moscow Engineering Physics Institute (Russia) and his M.S. degree in electrical engineering in 1997 from Worcester Polytechnic Institute (Worcester, MA). In 2002, he received his Ph.D. degree in electrical engineering from University of Washington (Seattle, WA), where he worked on modeling of arsenic deactivation and diffusion in silicon. In 2002, he joined the TCAD group at AMD in Sunnyvale, CA. His current interests include process and device modeling for high speed CMOS.

**Zudian Qin** received both the B.S. and M.S. degrees in physics from Peking University (Beijing, PRC), in 1990 and 1993, respectively. He worked as an electrical engineer for several years at Shenzhen University (PRC). Since 2000, he has been pursuing his Ph.D. in electrical engineering at the University of Washing-

ton in Seattle. He is currently working on the development of a kinetic lattice Monte Carlo (KLMC) simulator for the study of dopant diffusion and activation in silicon for future nanoscale devices.

**Milan Diebel** was born in Frankfurt am Main (Germany) in 1973. He received his Diplom-Physiker degree in theoretical nuclear physics from the Justus-Liebig-University Giessen (Germany) in 1999. He continued his education at the University of Washington (Seattle, WA) in the Department of Physics and received a M.S. in physics in 2000. Currently he is working on his Ph.D. in physics. His current research focuses on bridging the gap between atomistic and macroscopic scales, using density functional theory to predict macroscopic diffusion phenomena.