Interactions of B dopant atoms and Si interstitials with SiO$_2$ films during annealing for ultra-shallow junction formation

P. Kohli[a)]
Silicon Technology Development, Texas Instruments, Dallas, Texas 75432

A. Jain, S. Chakravarthi, and H. Bu
Silicon Technology Development, Texas Instruments, Dallas, Texas 75432

S. T. Dunham
Department of Electrical Engineering, University of Washington, Seattle, Washington 98195

S. Banerjee
Microelectronics Research Center, University of Texas, Austin, Texas 78758

In this work we present an investigation of the effect of oxide thickness on annealed B diffusion profiles. Experiments were specifically designed to determine the effect of varying oxide thickness on the B diffusion profile after annealing. Boron was implanted through a 50 Å screen oxide. Implant oxide was etched to varying degrees on different samples resulting in screen oxide thickness from 0 to 50 Å. On samples where the screen oxide was completely etched away, cap oxide was deposited with thickness varying from 0 to 50 Å. The implanted wafers were then spike annealed at 1050 °C. We found that the thicker the oxide during annealing, the deeper the B diffusion profile. A model of the Si–SiO$_2$ system based on the interactions of B dopant atoms and silicon interstitials with SiO$_2$ films is proposed to explain the experimental observations. The model takes into account the segregation of Si interstitials at the Si/SiO$_2$ interface and the diffusion of that Si in the oxide. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1884246]

I. INTRODUCTION

In the semiconductor industry, the dominant device used today is the Si-based metal oxide semiconductor field effect transistor (MOSFET). Improvements in the computer industry over the past 30 years have relied heavily on the ability to increase the speed of the Si MOSFET through the downward scaling of all vertical and lateral dimensions of the transistor. The scaling of the device dimensions not only leads to faster devices but also allows larger and more complex circuits to be implemented in a smaller area.[1] In the last decade, in order to continue conventional scaling of the source/drain junctions, the semiconductor industry has relied heavily on decreasing the implant energy, and also on minimizing the thermal budget of the activation anneal. Decreasing the implant energy puts the excess Si interstitials closer to the surface. The surface acts as an efficient sink for the interstitials, thus reducing transient enhanced diffusion (TED), which results from the interaction of excess Si interstitials with the dopant atoms.[2] For implant energies below 1 keV, TED can be nearly eliminated. Increasing the ramp rates of post-implantation anneals has also greatly reduced TED effects. With TED less pronounced for low implant energies and sharper anneal temperature profiles, surface reactions and related processes have again started to dominate the formation of ultra-shallow junctions. Interactions of dopant atoms and point defects with surface films and interfaces are becoming of paramount importance in determining the concentrations of dopants and point defects, and therefore the resulting diffusion profiles. Since interactions of point defects and dopant atoms play a central role in the integrated circuit fabrication processes, we present experiments that were designed to gain a fundamental understanding of the interactions of dopant atoms and point defects with thin oxide films.

II. EXPERIMENT

Typically, for source/drain extension formation, dopants have to be implanted through an oxide, called a screen oxide. The alternative of gate re-oxidation to repair the gate edge reaction ion etch (RIE) damage after the implant results in unacceptable oxidation-enhanced diffusion (OED).[3] Other motivations for implanting through a screen oxide include avoiding channeling, avoiding contamination from the implanter, and creating shallower implanted profiles without having to resort to ultra-low energy implants and associated energy contamination issues.

In a typical complementary metal-oxide semiconductor (CMOS) device flow this "implanted oxide" (since dopants have been implanted through it) will see a number of cleaning process steps in order to achieve low particle counts and metallic impurity levels. These cleans can result in etching away varying amounts of the implanted screen oxide. Table I summarizes the effect of two typical cleans used in the industry on the implanted screen oxide thickness, as measured by transmission electron microscopy. These cleans involve solution of NH$_4$OH, H$_2$O$_2$ and de-ionized water at room temperature (Clean1) and at 65 °C (Clean2). It should be noted that these cleans are calibrated routinely in the industry to ensure very limited etching of thermally grown oxide. How-

a)Electronic mail: p-kohli1@ti.com

97, 073520-1

TABLE I. The effect of two different cleans on the thickness of the implanted screen oxide.

Clean details	TEM oxide thickness (Å)
No clean (as-implanted)	57
Clean1	13
Clean2	0

FIG. 1. Boron SIMS profiles in Si showing the effect of cap oxide thickness on B diffusion profile. B was implanted through a 50 Å screen oxide and then etched away completely. Different thicknesses of cap oxide were deposited. Thicker oxide during annealing gives deeper junction.

ever, as shown in Table I, the same etches can cause substantial etching of implanted oxides. A B dose of 1.2×10^{15} atoms/cm^2 was implanted through a 50 Å (target thickness) thermally grown oxide for all these samples. It should be noted that the initial screen oxide thickness as determined by transmission electron microscopy (TEM) analysis was 57 Å and it is reduced to 13 and 0 Å for Clean1 and Clean2, respectively. In our experiments, we have tried to imitate the effect of cleans by varying the implanted screen oxide thickness prior to annealing by using well-tailored etches. Therefore, besides giving a fundamental understanding of the effect of varying oxide thickness on the dopant diffusion profile, it also makes the experiments very relevant to the industry.

The starting material was *n*-type silicon with $\langle 100 \rangle$ crystal orientation. The wafer splits are summarized in Table II. A boron dose of 1.2×10^{15} atoms/cm^2 was implanted through the 50 Å thermally grown screen oxide at 1.3 keV, corresponding to a standard source/drain extension implant on all wafers. In a subset, labeled Set A, after B implantation, the thermally grown screen oxide was etched back to different thicknesses and then the wafers were annealed. For reference, one wafer in this subset was preserved without any screen oxide etch (50 Å oxide intact). All but one wafer from this subset were spike annealed at 1050 °C. In another subset, labeled Set B, after B implantation, the screen oxide was etched off completely and two different thicknesses of rapid thermal chemical vapor deposition (RTCVD) cap oxide (30 or 50 Å) were deposited. No oxide was deposited on one wafer after etching the screen oxide. All wafers in this subset were also spike annealed at 1050 °C. All anneals were performed in a highly-inert N$_2$ ambient with very low levels of oxygen ($<$100 ppm) to avoid any oxidation enhanced diffusion.[4]

B profiles in Si were obtained using secondary-ion-mass spectroscopy (SIMS). All samples were dipped in hydrofluoric (HF) acid before SIMS analysis to remove any surface oxide. SIMS measurements were performed using a CAMECA IMS 6f magnetic sector instrument. All measurements were performed using an O$_2^+$ primary beam and detecting positive secondary ions. A primary oxygen beam with an impact energy of 800 eV was used. The angle of incidence was approximately 42°. The beam current and raster size were adequate to provide about 0.5 Å/s erosion rate. An oxygen backfill was applied to ensure the surface was fully oxidized during depth profiling. Some of the SIMS analyses were repeated in order to confirm the results. Sheet resistance values were obtained using a four-point probe.

III. RESULTS

A. Effect of RTCVD cap oxide thickness on B diffusion

Figure 1 shows the B SIMS profiles after a 1050 °C spike anneal for different thicknesses of the RTCVD cap oxide during annealing. All of these samples had a B implant (Set AB). The as-implanted profile is shown for comparison. The sample with the thickest RTCVD cap oxide (50 Å) results in the deepest junction depth. The sample with no oxide results in the shallowest profile, the 30 Å RTCVD cap oxide gives an intermediate junction depth. Table III shows the sheet resistance values for these samples.

B. Effect of screen oxide thickness on B diffusion

Figure 2 shows the B SIMS profiles after a 1050 °C spike anneal for different thicknesses of the screen oxide

TABLE II. Summary of wafer-processing conditions.

Process steps/wafers	Set A				Set B		
Screen (thermal) oxide 50 Å	X	X	X	X	X	X	X
B implant	X	X	X	X	X	X	X
Etch screen oxide to 30 Å		X					
Etch screen oxide to 15 Å			X				
Etch screen oxide completely (0 Å)					X	X	X
Cap (RTCVD) oxide 30 Å						X	
Cap (RTCVD) oxide 50 Å							X
1050 °C spike	X	X	X		X	X	X

TABLE III. Effect of oxide thickness—cap and screen—on sheet resistance. The trend corroborates SIMS profiles as shown in Figs. 1 and 2.

Screen oxide		Cap oxide	
Oxide thickness (A)	Sheet resistance (Ohms/sq)	Oxide thickness (A)	Sheet resistance (Ohms/sq)
15	331	0	397
30	308	30	397
50	302	50	364
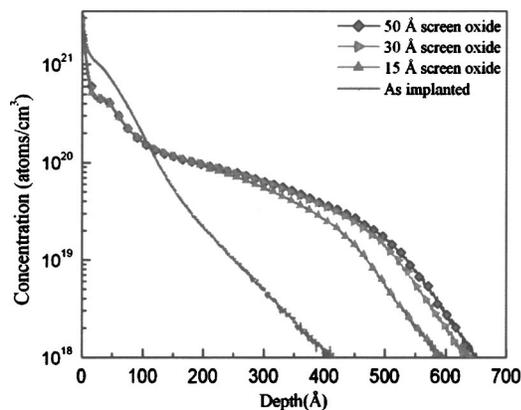
FIG. 2. Boron SIMS profiles in Si showing the effect of screen oxide thickness on B diffusion profile. B was implanted through a 50 Å screen oxide and then etched to different thicknesses. Thicker oxide during annealing gives deeper junction.

during annealing (Set AA). All these samples had a B implant. The as-implanted profile is shown for comparison. As with deposited cap oxides, thicker oxides result in the deeper junction depths. Table III shows the sheet resistance values for these samples.

IV. DISCUSSION

It is important to note that by etching the oxide to different thicknesses or depositing different thicknesses of RTCVD cap oxide after the implant, all samples have identical initial dopant and damage distributions in the Si. As is evident from Figs. 1 and 2, the samples with the thickest oxide result in the deepest junctions and the samples with the thinnest oxide result in the shallowest junctions. This trend is observed for both thermally grown screen oxides and RTCVD deposited cap oxides. The sheet resistance measurements shown in Table III further confirm the observed trend. The sample with the deepest profile results in the lowest sheet resistance in each case as more of the dopant is active rather than clustered. Theoretical calculations of sheet resistance based on empirical formulas for the given SIMS profiles agree well with the experimental measurements.[5]

Kasnavi *et al.* suggested that for B implants dose loss is due to segregation of B into the bulk of the oxide.[6] Their model predicts that a thicker oxide would result in more bulk segregation and therefore more dose loss. Increased dose loss could only result in shallower (or similar, if TED was dominant) diffusion profiles. However, we clearly see that our results for different oxide thicknesses (for both cap and screen oxides) show deeper profiles for the samples with the thicker oxides. Therefore bulk segregation and resultant B dose loss cannot explain the deeper junctions with thicker oxides.

Another possible explanation could be that the outdiffusion flux of B from the oxide into the ambient is the dominant flux, and therefore the thicker the oxide the less the outdiffusion flux and hence the deeper the profile. However, it needs to be noted that the integrated dose values for the diffused SIMS profiles show similar values for the profiles with different oxide thicknesses. Based on reports of B dose
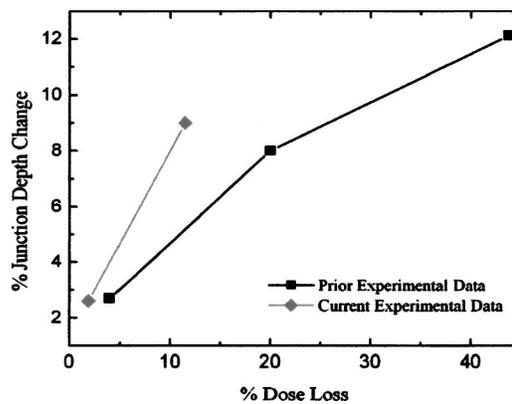
FIG. 3. Percentage dose loss vs. percentage change in junction depth for current experiments and prior dose loss related experiments. Clearly the current experimental data does not agree with the prior dose loss data.

loss for different nitride spacer chemistries,[7] experimentally observed percentage changes in junction depth associated with varying amounts of dose loss for similar B doses and anneals are shown in Fig. 3.[7] Also shown in this figure is the change in junction depth versus dose loss value for the current work. Clearly our data shows substantially less dose loss for the same change in junction depth compared with prior experimental observations for which dose loss was identified as the primary cause of junction shift. Also, in an attempt to simulate the observed junction depth differences based on B dose loss, we set the outdiffusion flux of B from oxide into the ambient to a high value and considered B flux in the oxide as the rate-limiting flux. We found that we could match the experimentally observed differences in the junction depth with different oxide thicknesses by adjusting the B dose loss from the Si into the oxide and then further into the ambient, but doing so resulted in perceptibly different profiles throughout the silicon. In contrast, the experimental profiles appear identical near the surface for different oxide thicknesses and the only differences are seen towards the tail of the diffused profiles. This suggests that B dose loss due to outdiffusion is not the reason for the observed differences in the diffused profiles for different oxide thicknesses.

High temperature SiO_2 decomposition at the oxide/Si interface has been reported in the literature. At high temperatures and low oxygen partial pressure, SiO_2 decomposition can occur according to the reaction $Si+SiO_2 -> SiO$, where the Si substrate acts as a source of the Si for the reaction. The process can be extremely nonuniform, resulting in the nucleation and growth of voids with little overall thinning of the oxide.[8–12] The decomposition of SiO_2 into SiO requires the net diffusion of Si within the oxide (as Si or SiO or counterflux of O vacancies) and the formation of SiO which evaporates at the SiO_2/gas interface. Based on observations of dopant diffusion and stacking fault shrinkage, Ahn[13] suggested that vacancy generation at the SiO_2/Si interface resulted from SiO formation in thin oxide films in inert ambients at high temperatures. He found that P diffusion was retarded, as diffusion was enhanced, and the shrinkage rate of preexisting stacking faults was increased. These observations strongly suggest that a vacancy supersaturation and self-interstitial undersaturation exist under thin SiO_2 layers during annealing in Ar.[13]
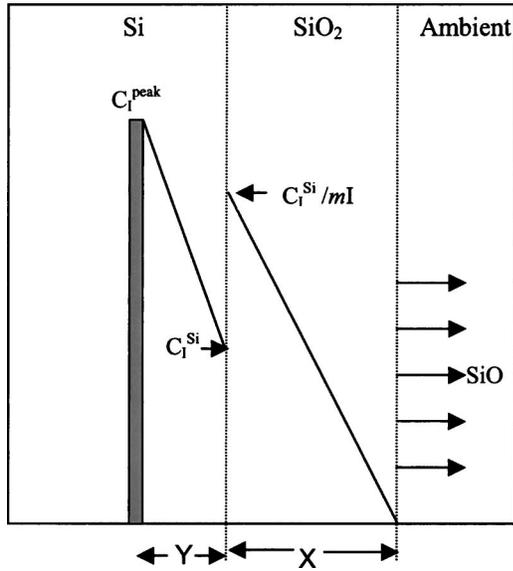
FIG. 4. Schematic illustration of the concentrations of reacting/diffusing species during inert ambient annealing of thin oxide films.

Dunham[2] modeled the Si/SiO_2 system more quantitatively by considering a proposed set of processes: desorption of SiO into the gas phase, diffusion of Si across the oxide, segregation of Si interstitials at the Si/SiO_2 interface, interface recombination, and diffusion of interstitials in the substrate. In his model, he assumed that a constant segregation coefficient m_I relates the ratio of the concentration of Si interstitials in the Si to that in the oxide, and that there are no reactions involving interstitials in the oxide. Recent calculations using density functional theory show the ease of diffusion of interstitials from Si to the Si/SiO_2 interface and discuss the migration of Si interstitials into the SiO_2.[14] With his model he was able to explain a broad range of data both under oxidizing and nonoxidizing conditions and also estimate the amount of thinning that would be consistent with the interstitial undersaturation observed by Ahn. Dunham's model suggests a strong dependence on oxide thickness and that thickness dependence is matched well by a model assuming the diffusion of a Si species across the oxide.[2]

We believe that the diffusion observed in our samples as a result of the 1050 °C spike has an initial TED component followed by an equilibrium component. The effect of oxide thickness variation on the equilibrium B diffusion can be understood by Dunham's model. It predicts retarded B equilibrium diffusion in the presence of thin oxide on the surface because of interstitial undersaturation. In order to understand the effect of oxide thickness variation on the TED component of B diffusion, we use a similar analysis under conditions of interstitial supersaturation (due to implantation). Figure 4 illustrates schematically the concentration of active species during inert ambient annealing for a Si/SiO_2 system under conditions of interstitial supersaturation. The effect of oxide thickness can be understood by a simple analysis of TED. It has been experimentally observed that the B diffusivity enhancement during TED is nearly independent of the B implant damage for initial stages and after some time period τ the enhancement goes away.[15,16] This would imply

that the excess interstitial concentration is approximately fixed during the early stages of TED and after some time it comes down to the equilibrium value. Reports by Eaglesham *et al.* revealed {311} defects as the source of the interstitials during TED.[17]

Let C_I^{peak} be the effective interstitial solubility associated with the formation of {311} defects. Let Y be the average depth into Si from the Si/SiO_2 interface of net interstitial distribution. This depth Y could be the projected implant range for a non-amorphizing implant or the end-of-range depth for an amorphizing implant. The total dose of excess interstitials (after initial recombination) is Q. Based on a "+1" damage model, Q is equal to the implanted dose.[6] We assume that after the initial recombination of interstitials and vacancies (starting with a +1 interstitial dose), there are no interstitial reactions with vacancies and lattice defects in the Si and the dominant flux of interstitials is towards the surface. To simplify our analysis further, we assume that the concentration of the mobile Si species at the $SiO_2/ambient$ interface is zero. Now we can assume a steady state flux balance at the SiO_2/Si interface between the flow of Si interstitials from the substrate to the interface, recombination at the interface and flow across oxide,

$$[D_I^{Si}\{C_I^{peak} - C_I^{Si}\}]/Y = [\sigma\{C_I^{Si} - C_I^*\}]$$
$$+ [D_I^{SiO_2}/(X\{C_I^{Si}/m_I\})] \qquad (1)$$

where C_I^* is the interstitial concentration under equilibrium conditions, X is the oxide thickness, σ is the interface recombination velocity for interstitials, D_I^{Si} is the diffusivity of the interstitials in Si, $D_I^{SiO_2}$ is the diffusivity of excess Si in SiO_2, and m_I is the segregation coefficient for excess Si between Si and oxide. Furthermore, we can estimate the TED time, $\tau(=dose/flux)$, as the time taken for all the excess interstitials to escape out of the system,

$$\tau = Q/[(D_I^{Si}\{C_I^{peak} - C_I^{Si}\})/Y]$$
$$= Q/([\sigma\{C_I^{Si} - C_I^*\}] + [D_I^{SiO_2}/(X\{C_I^{Si}/m_I\})]). \qquad (2)$$

Ignoring the recombination flux (negligible interface recombination when implanted dose is much higher than the number of interface traps),

$$\tau = Q/[D_I^{SiO_2}/(X\{C_I^{Si}/m_I\})]. \qquad (3)$$

During TED, the interstitial supersaturation is C_I^{peak}/C_I^*, so the amount of excess diffusion expected during TED is given by,

$$\sqrt{(Dt)}_{TED} = \sqrt{(D_B^* f_I [C_I^{peak}/C_I^*]\tau)}$$
$$= \sqrt{(D_B^* f_I [C_I^{peak}/C_I^*](Q/[D_I^{SiO_2}/(X\{C_I^{Si}/m_I\})]))}$$
$$= \sqrt{(D_B^* f_I [C_I^{peak}/C_I^*] \cdot [Q/D_I^{SiO_2}]X\{C_I^{Si}/m_I\})} \qquad (4)$$

where D_B^* is the B diffusivity under equilibrium conditions and f_I is the fraction of B diffusion associated with interstitial mechanism. This formalism clearly shows that the thicker the oxide (X), the greater the B diffusion. This model therefore explains why we would observe a deeper junction for a thicker oxide.
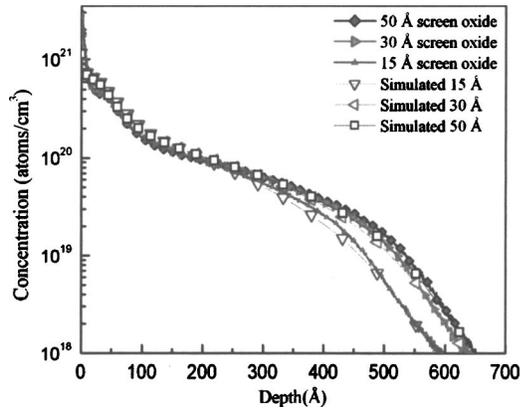
FIG. 5. Comparison between model and data for diffusion of B under different screen oxide thicknesses during annealing in an inert ambient.

For a more precise analysis we ran simulations considering transient diffusion/recombination in the silicon, the as-implanted B SIMS profile and a "+1" damage model. However, one of the major barriers to such modeling is that the reported values for parameters such as interstitial diffusivity, Si atom diffusivity in oxide, and interstitial segregation coefficient between Si/SiO$_2$ vary over several orders of magnitude.[2] The effect of the decomposition of thin oxide films can be captured by understanding that the phenomenon drives the interstitial concentration at the interface to values below the equilibrium concentrations. In our simulations we fixed C_I/C_I^* at the interface to different values in order to match the profiles. For thin films, the Si flux in the oxide being very high could result in high degrees of interstitial undersaturation. Figure 5 shows the simulated profiles and SIMS profiles for the different screen oxide thicknesses. Simulated B diffusion profiles show good agreement with the SIMS profiles. The C_I/C_I^* values at the interface for the simulations for different oxide thicknesses are shown in Fig. 6. In our case it should be noted that the model was calibrated assuming that the 50 Å screen oxide results in no interstitial undersaturation at the interface. This assumption makes it possible for us to quantify the thickness effect on the C_I/C_I^* undersaturation (as shown in Fig. 6). It seems to be a reasonable assumption considering that the differences
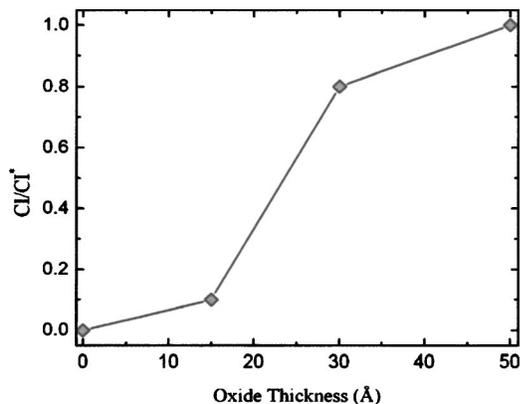
in junction depth diminish going from 30 to 50 Å oxide (as compared to the differences observed between 15 and 30 Å screen oxide profiles).

Comparing the B diffusion profiles for samples with 50 Å screen oxide and 50 Å RTCVD cap oxide, it is evident that the RTCVD cap oxide results in a higher dose loss from the Si than the screen oxide during annealing. The integrated dose numbers are as different as 5.4×10^{14}/cm^2 and 7.8×10^{14} cm^2 for the RTCVD cap oxide and screen oxide samples, respectively. The same trend is observed between the RTCVD oxide and screen oxide for the other thicknesses as well. It is expected that the RTCVD oxide will result in a higher dose loss into the oxide since the deposited oxide has no B to start with while the screen oxide receives substantial B dose during implantation. So for the RTCVD cap oxide it could be expected that more B would be lost due to segregation into the oxide. However, our simulations show that the differences that could be expected based on segregation are much smaller than those observed experimentally. Kohli *et al.* have reported the presence of high levels of H in the as-deposited oxides.[7] Many reports in the literature have discussed enhanced B diffusion in the oxides due to the presence of H.[18–20] While it is believed that the H could diffuse out of the deposited oxides into the ambient upon annealing, it might still result in highly enhanced B diffusion in the initial phase of annealing.

It is interesting to note in Fig. 1 that the profile with no oxide results in the shallowest junction. However, the sheet resistance value is the same as that of the sample with 30 Å cap despite a shallower X_j. Looking carefully at the profile, it can be seen that the concentration of B is higher (between 100 and 170 Å into Si) for the sample with no oxide as compared to samples with the cap. This would explain why the sheet resistance values are similar between the 30 Å cap oxide and no oxide cases even though the no oxide case has a lower X_J. The shallower diffusion profile indicates that the bare Si serves as a much more effective interstitial sink than the RTCVD cap oxide/Si interface. This understanding could provide ways to form shallower junctions without compromising the sheet resistance.

V. CONCLUSION

In the last decade, in order to continue conventional scaling of the source/drain junctions, the semiconductor industry has relied heavily on decreasing the implant energy, and also on minimizing the thermal budget of the activation anneal. However, with transient enhanced diffusion (TED) being less pronounced for low implant energies and sharper anneal temperature profiles, surface reactions and related processes are starting to dominate the formation of ultra-shallow junctions. We have studied the effect of surface oxide layer on B diffusion profile. We have found that thinner oxides result in shallower junctions. We have successfully modeled the effect of oxide thickness on junction depth. The model takes into account the segregation of Si interstitial at the Si/SiO$_2$ interface and diffusion of that Si in the oxide.

FIG. 6. Interstitial undersaturation vs oxide thickness—based on simulations shown in Fig. 5.

[1] P. Packan, MRS Bull. **25**, 18 (2000).

[2] S. T. Dunham, J. Appl. Phys. **71**, 2, 1992.

[3] R. Lindsay, *Proceedings of the Sixth International Workshop on Fabrication, Characterization, and Modeling of Ultra Shallow Doping Profiles in Semiconductors (USJ-2001)*, Napa, California, USA, 22–26 April, 2001, pp. 255.

[4] M. Miyake, J. Appl. Phys. **57**, 6, 1985.

[5] W. Johnson, Solid-State Electron. **13**, 951, 1969.

[6] R. Kasnavi, P. B. Griffin, and J. D. Plummer, Mater. Res. Soc. Symp. Proc. 610, 2000.

[7] P. Kohli, A. Jain, H. Bu, S. Chakravarthi, C. Machala, S. T. Dunham, and S. K. Banerjee, J. Vac. Sci. Technol. B **22**, 471, Jan 2004.

[8] R. Tromp, G. W. Rubloff, P. Balk, F. K. Legoues, and E. J. Vanloenen, Phys. Rev. Lett. **55**, 2332, 1985.

[9] K. Hofmann, G. W. Rubloff, and R. A. McCorkle, Appl. Phys. Lett. **49**, 1525, 1986.

[10] K. Hofmann and S. I. Raider, J. Electrochem. Soc. **134**, 240, 1987.

[11] G. W. Rubloff, K. Hofmann, M. Liehr, and D. R. Young, Phys. Rev. Lett. **58**, 2379, 1987.

[12] M. Liehr, J. E. Lewis, and G. W. Rubloff, J. Vac. Sci. Technol. A **5**, 1559, 1987.

[13] S. T. Ahn, Ph.D. thesis, Stanford University, 1988.

[14] T. Kirichenko, Ph.D. thesis, University of Texas at Austin, 2005 (to be published).

[15] R. Angelucci, F. Cembali, P. Negrini, M. Servidori, and S. Solmi, J. Electrochem. Soc. **134**, 3130, 1987.

[16] P. A. Packan, Ph.D. thesis, Stanford University, 1991.

[17] D. J. Eaglesham, T. E. Haynes, H. J. Grossmann, D. C. Jacobson, P. A. Stolk, and J. M. Poate, Appl. Phys. Lett. **70**, 3281, 1997.

[18] R. Fair, IEEE Electron Device Lett. **17**, 11, Nov. 1996.

[19] Y. Shacham-Diamand and W. G. Oldham, J. Electron. Mater. **15**, 229, 1986.

[20] T. Aoyama, K. Suzuki, H. Tashiro, Y. Tada, Y. Kataoka, H. Arimoto, and K. Horiuchi, Jpn. J. Appl. Phys., Part 1 **38**, 2381, 1999.