Journal of The Electrochemical Society, **150** (12) G758-G765 (2003)
0013-4651/2003/150(12)/G758/8/$7.00 © The Electrochemical Society, Inc.

Parameter Sensitivity Analysis Applied to Modeling Transient Enhanced Diffusion and Activation of Boron in Silicon

R. Gunawan, M. Y. L. Jung, R. D. Braatz, and E. G. Seebauer[z]

Department of Chemical Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

Transient enhanced diffusion (TED) during annealing of implanted boron in silicon greatly impedes the formation of junctions that are sufficiently shallow for advanced complementary metal oxide semiconductor devices. For reasons of cost and efficiency in process development, detailed modeling of TED is often used for designing annealing procedures. However, because model validation depends primarily on fitting experimental dopant profiles, and because realistic models contain dozens of parameters that are poorly known, development of a unique set of rate parameters with true predictive capability has proven elusive. Here we have employed formal parameter sensitivity analysis by the finite difference method to show that the activation energies most critical to know accurately are those for interstitial boron diffusion, kick-in, and dissociation of the $(B_s - Si_i)^+$ complex to liberate either interstitial B (kick-out) or Si. Maximum likelihood estimation is also applied to the literature for interstitial cluster formation to determine a most likely set of activation energies for cluster dissociation.
© 2003 The Electrochemical Society. [DOI: 10.1149/1.1619992] All rights reserved.

Manuscript submitted October 17, 2002; revised manuscript received May 5, 2003. Available electronically October 10, 2003.

Forming ultrashallow junctions in Si-based microelectronic logic devices is becoming increasingly critical as device dimensions continue to shrink. Advanced complementary metal oxide semiconductor devices will require junction depths between 13 and 22 nm in the source and drain extension regions by the year 2005.[1] Current technology for junction formation relies almost exclusively on ion implantation to introduce dopants into the substrate. Although junction depths can be made shallower by reducing the implant energy, the effectiveness of this approach has been limited by the need to anneal the resulting structure to activate the dopant electrically and to eliminate implant-induced defects in the crystal structure. These defects mediate transient enhanced diffusion (TED) of the dopant, often leading to significant increase in junction depth.

For reasons of cost and efficiency in process development, detailed modeling of TED has attracted increasing attention for designing suitable annealing procedures.[2,3] However, the state of current modeling is far from satisfactory. One reason is that many elementary kinetic steps contribute to the experimental observable, typically a dopant depth profile obtained by secondary ion mass spectroscopy (SIMS). Hence, all models include numerous rate parameters, such as activation energies and pre-exponential factors for diffusion. Independent determinations of the parameters have been difficult to make.[4] In practice, therefore, many parameter sets have been developed primarily according to their ability to fit experimental SIMS profiles. The large number of parameters of course provides many degrees of freedom for fitting, impeding the ability to develop a unique set. Because simulated profiles are often compared only to a small body of experimental data, the predictive capability of most TED models outside their tested range is therefore subject to serious doubt. Worse, despite the many degrees of freedom available, adequately matching experimental data has been difficult without resorting to uncomfortably ad hoc approaches.[5]

In light of these problems, firmly grounded procedures are needed for estimating the most likely values of kinetic rate parameters. Elsewhere, we have discussed how to use maximum likelihood (ML) estimation in conjunction with established literature for parameter estimation, together with multivariate statistics to quantify accuracy.[4] Parameter sensitivity analysis based on direct simulation of experimental profiles offers an alternate tool with similar mathematical rigor to investigate the behavior of a process model.

The theory behind sensitivity analysis has been well developed[6,7] and applied in many fields of science and engineering.[8] In such analysis, "sensitivity coefficients" describe the variations in the outputs of a process model in response to perturbations in the parameters. The analysis suggests which parameters need to be measured or calculated most accurately by independent means, and which can be largely ignored. Sensitivity coefficients can also offer significant physical insights into the key elementary steps governing the overall reaction-diffusion network.

In this work, we apply sensitivity analysis by the finite difference method to a TED model recently developed in this laboratory.[4] Sensitivity coefficients are estimated using simulations of the TED model implemented in the profile simulator FLOOPS 2000.[a,9] Implementation of the analysis requires prior development of a set of activation energies for interstitial cluster dissociation, which we develop by application of ML estimation to the available literature. The results show that the activation energies most critical to know accurately are those for interstitial boron diffusion, kick-in, and dissociation of the $(B_s - Si_i)^+$ complex to liberate either interstitial B (kick-out) or Si.

Theory of Parameter Sensitivity Analysis

The matrix $\mathbf{S}$ of sensitivity coefficients describes quantitatively the change in process behavior in response to variations in the governing rate parameters. Process behavior in TED can be embodied by common metrics such as junction depth and dose loss as well as the levels of dopant activation. Let $\mathbf{y}$ denote the vector of the various process metrics chosen, and $\boldsymbol{\theta}$ denote the vector containing all rate parameters. The matrix $\mathbf{S}$ of sensitivity coefficients local to the region described by $\boldsymbol{\theta}$ can be described mathematically as

$$\mathbf{S}_{i,j} = S(\mathbf{y}_i;\boldsymbol{\theta}_j) = \frac{\partial \mathbf{y}_i(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \quad [1]$$

where the subscripts refer to the elements of the vector or the matrix.

There are three methods for evaluating the elements $\mathbf{S}_{ij}$: the direct differential method, the Green's function method, and the finite difference method.[8] The direct differential method utilizes the differential equations for the process model to derive differential equations describing the sensitivity coefficients. The sensitivity equations are obtained by taking the derivatives of the differential equations for the process with respect to each parameter. This approach is intuitively natural, but it requires solving $(m + 1) \times n$ equations simultaneously, where m is the number of parameters and n is the number of process outputs. The Green's function method seeks to reduce the number of equations by separating the problem into homogenous and nonhomogeneous parts and solving them

[a] By Mark E. Law of the University of Florida and Al Tasch of the University of Texas at Austin.

[z] E-mail: eseebaue@uiuc.edu

Table I. Activation energies for interstitial diffusion and cluster association.

Reaction	Symbol[a]	Activation energy (eV)	Reference	Method[b]
B_i^+ diffusion	E_{diff,B_i^+}	0.37 ± 0.04[c]	4	ML
Si_i^{+2} diffusion	$E_{diff,Si_i^{+2}}$	0.72 ± 0.03[c]	4	ML
$B_i^+ + Si_s \rightarrow (B_s - Si_i)^+$	E_{ki}	0.50 ± 0.1	4	ML
$B_m - Si_n + B_i^+ \rightarrow B_{m+1} - Si_n,\ n,m \geq 0$[d]	$E_{assoc,B}$	0.37 ± 0.04[c]	4	Assumed $= E_{diff,B_i^+}$
$B_m - Si_n + Si_i^{+2} \rightarrow B_m - Si_{n+1},\ n,m \geq 0$[d]	E_{assoc}	0.72 ± 0.03[e]	4	Assumed $= E_{diff,Si_i^{+2}}$

[a] For clarity in focus on cluster effects, the present work uses slightly different notation than related publications from this laboratory.[4] Where two symbols are given here, the latter symbol appears in Ref. 4.
[b] ML = Maximum likelihood estimation.
[c] The pre-exponential factor for this diffusion has been assumed to be 1×10^{-3} cm^2/s.[4]
[d] For $m \geq 1$, one of the boron atoms is presumed to be substitutional. Also, m and n must obey $m + n \geq 2$ and $m + n \leq 4$. Finally, if $n = 0$, then m is assumed to obey $m \leq 1$. (No pure boron clusters larger than dimers form.)
[e] The pre-exponential factor for this association reaction has been assumed to be 3×10^{-10} cm^2/s,[4] regardless of cluster size. Also, if two free Si interstitials recombine, the stochiometric factor of 2 in the rate constant for diffusion limitation (which includes $D_{Si} + D_{Si}$) has been neglected.

separately. This procedure requires solving $n \times n$ equations plus n integrals,[8] which offers significant computational advantages if m is much smaller than n. Nevertheless, both methods are applied only when the process model is cast in terms of ordinary differential equations.

Unlike the two previous approaches, the finite difference method completely avoids additional differential equations that need to be solved simultaneously. Instead, this method estimates the sensitivity coefficients using the finite difference method

$$S(\mathbf{y}_i;\boldsymbol{\theta}_j) \approx \frac{\Delta \mathbf{y}_i}{\Delta \boldsymbol{\theta}_j} = \frac{\mathbf{y}_i(t,\boldsymbol{\theta}_j + \Delta\boldsymbol{\theta}_j) - \mathbf{y}_i(t,\boldsymbol{\theta}_j - \Delta\boldsymbol{\theta}_j)}{2\Delta\boldsymbol{\theta}_j} \quad [2]$$

where $\Delta\boldsymbol{\theta}_j$ is a small variation of $\boldsymbol{\theta}_j$. In principle, this approach requires solution of n differential equations twice for each parameter. Note that for a transient set of equations such as those describing TED, $\mathbf{S}_{ij}$ can be time-dependent. The accuracy of the finite difference estimate depends on the magnitude of the variation $\Delta\boldsymbol{\theta}_j$. The value for $\Delta\boldsymbol{\theta}_j$ should be small enough that the effects of higher order terms are negligible, but large enough that the estimate is insensitive to simulation errors appearing in the numerator (*i.e.*, error in $\mathbf{y}_i$). The sensitivity coefficients for different dependent variables $\mathbf{y}_i$ sometimes vary greatly in magnitude. Hence, the parameter variation $\Delta\boldsymbol{\theta}_j$ must vary correspondingly. Here, 5, 10, and 20% variations for $\Delta\boldsymbol{\theta}_j$ are computed and compared. Note that, in general, sensitivity coefficients that are very small cannot be computed accurately due to the limited accuracy of any given simulation code. Such coefficients are simply reported as being less than some fixed value.

Overall TED Model

This simulator solves the coupled mass balance equations for interstitials, vacancies and clusters. These equations have the general form for species j

$$\frac{\partial N_j}{\partial t} = -\frac{\partial J_j}{\partial x} + G_j \quad [3]$$

where N_j denotes concentration and G_j a net generation rate. The flux J_j comprises terms due to both diffusion and drift in response to electric fields. The electric fields are obtained by solution of Poisson's equation. We implemented FLOOPS with the rate expressions and parameters shown in Table I together with no-flux surface boundary conditions for all species[b,10] and no surface Fermi level pinning. Concentrations of charged interstitial B and Si species were computed according to Fermi statistics as described elsewhere.[11] Initial conditions on the profiles for Si_i were set using the well-known "+1" model in which $[Si_i]$ exactly tracks the local concentration of implanted boron. For total boron, we employed experimental as-implanted profiles as initial conditions, with a fixed fraction of one-fifth of the total boron in substitutional sites, in accord with the suggestion of de la Rubia *et al.*[12] and Kobayashi *et al.*[13]

Figure 1 shows the heating program employed, which corresponded to that used for an experimental data set to which the simulations were compared. The experimental wafers, processed by International Sematech, were implanted with B at 0.60 keV with a fluence of 2×10^{15} ions/cm^2 at 0° tilt. Results for a comparable set of experiments have also been reported by Downey *et al.*[14] The heating rate β was 150°C/s, leading to a junction depth X_j of 54 nm. Similar X_j results were obtained experimentally by Downey *et al.*[14] for roughly equivalent implant and rapid thermal annealing (RTA) conditions.

Cluster Model: Maximum Likelihood Estimation

Cluster sizes.—During implantation and annealing, it is well known that clusters of atoms form and during thermal annealing can absorb and release the free interstitial atoms that are primarily

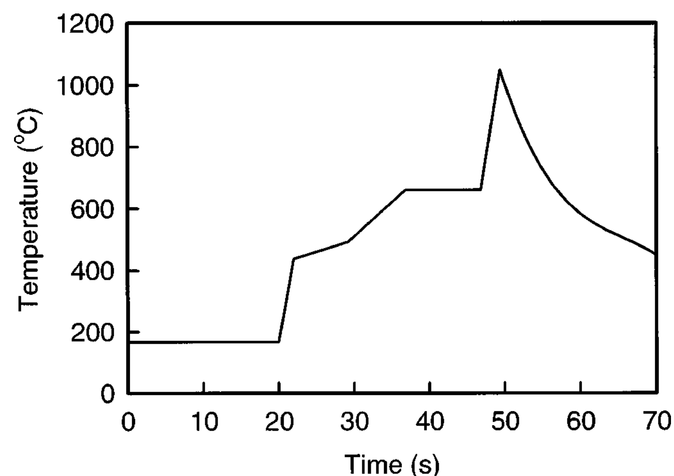

Figure 1. Experimental heating profile.

[b] No-flux boundary conditions represent an approximation that suffer deficiencies as discussed in Ref. 10. However, in the regime of the simulations, the conclusions of this paper are insensitive to the details of these conditions.

responsible for TED. There is considerable evidence that these clusters can comprise pure B, pure Si, or both. Due to the small number of differential equations that can be tracked by FLOOPS 2000, cluster sizes involving interstitials were limited to five atoms for pure Si clusters. We doubt that this restriction imposes serious limitations on simulation accuracy because experimental observations from spike RTA of sub-keV implanted boron showed no evidence for the formation of large clusters.[15] Regarding boron, there is no evidence for pure B clusters larger than two atoms that involve interstitial atoms. Instead, B interstitials seem to be more effectively captured by mixed B-Si clusters.[16,17] Hence, cluster sizes were limited to two atoms for pure B and five for mixed B-Si. The initial concentrations of all clusters and complexes were set to zero.

Cluster growth.—The association reaction between Si_i or B_i and a cluster is second order. Although an activation barrier may exist in principle when these species get close enough to react, the negative free energy of formation for the complex gives reason to believe that the complex forms with no barrier. In accord with much literature in the field, a rate expression describing standard diffusion limitation by reactants[18] seems warranted

$$r_{\mathrm{assoc}} = k_{\mathrm{assoc}}[\mathrm{cluster}][\mathrm{B_i\ \ or\ \ Si_i}] \quad [4]$$

with

$$r_{\mathrm{assoc}} = 4\pi a D_{\mathrm{B_i}} \ \text{ or } \ r_{\mathrm{assoc}} = 4\pi a D_{\mathrm{Si_i}} \quad [5]$$

where clusters have been assumed to be much less mobile than free interstitials. Here, a represents a reaction distance or capture radius.

There exist significant questions about what value a should take. The capture radius may depend on the identities of the atoms and clusters involved as well as their charge states and orientation within the Si lattice.[19] For example, Coulombic attraction between species of opposite charge probably lead to increased values of a, while like charges lead to decreased values. Effects of this sort remain inadequately treated in the literature, so that even for association of single atoms, a has been assigned a variety of values, ranging from the nearest neighbor distance of 2.7 Å[20,21] to the Si lattice constant[22] (~5 Å) to even larger values near 7 Å.[19] Because literature reports of a represent mostly assumptions rather than measured values, and because a varies with reaction stoichiometry and other factors, quantitative methods employed here for assigning a most probable value are not suitable. Instead, we treat it as a fixed parameter with an assumed value of the Si nearest neighbor distance, 2.7 Å. Although other choices are certainly defensible, they are unlikely to vary from this one by more than a factor of about two. This modest range does not affect the primary conclusions to be drawn below, although it may affect some of the precise details for the activation energies demonstrating the largest parameter sensitivity.

Cluster dissociation.—Dissociation kinetics are more problematic, because the rate expressions used in the literature up to now have inadequacies, and there is no agreement in the literature regarding the magnitudes of the dissociation energies. The rate expression commonly employed in literature for dissociation kinetics assumes the following form[23]

$$r_{\mathrm{dis}} = \frac{D_{\mathrm{int}}}{a^2}\exp(-E_{\mathrm{b}}/kT)[\mathrm{cluster}] \quad [6]$$

where E_b denotes the binding energy and D_{int} the interstitial diffusivity. Equation 6 is problematic in several ways, but its greatest flaw originates from its assumption that dissociation involves two sequential steps: interstitial release from the cluster followed by a diffusional hop away from the cluster. However, except for $(B_s{-}Si_i)^+$, the barrier to interstitial release (for which the binding energy sets only a lower bound) is significantly larger than that for diffusional hopping. Because the pre-exponential factor is likely to be the same for both release and hopping (*i.e.*, close to a Debye frequency as discussed elsewhere[4]), interstitial release must be rate limiting in this reaction sequence. It can easily be shown that in a reaction sequence including a rate-limiting step, steps after the rate limiting one exert no influence on the overall rate. Thus, rate expressions that include parameters from steps following the rate limiting one must, in principle, be incorrect.

In general, simple models for elementary dissociation reactions employ first-order kinetics and cast rate constants as the mathematical product of an attempt frequency and an exponential Boltzmann factor containing a transition-state activation barrier (which may not be the same as the binding energy). We employ that picture here. Regarding the pre-exponential factor, we chose the Debye frequency for Si of about $6 \times 10^{12}\ \mathrm{s^{-1}}$, subject to various caveats discussed elsewhere.[4]

The literature reports several values for dissociation energies, which are believed to depend upon cluster size. For example, Pelaz *et al.*[24] suggested a square-root dependence on size for pure Si clusters. Cowern *et al.*[25] reported a more complicated dependence having minima at specific sizes, most likely near four and eight atoms. The experimental literature is augmented by a density functional theory (DFT) calculations, especially for dimers and large clusters. With this fairly rich literature base, we applied ML estimation to these species (aside from $(B_s - Si_i)^+$, which we have treated by this method previously[4]). The details of the parameter estimation method have already been reported elsewhere.[4]

Note, however, that DFT calculations should be employed only with many caveats. Most DFT work reports thermodynamic formation energies, which involve the initial and final states, rather than true activation barriers, which involve the initial and transition states. In general, barriers for endothermic reactions like cluster dissociation often match thermodynamic formation energies, but this need not always be the case. Thermodynamic energies set only a minimum value for activation barriers. Note also that most quantum calculations are valid only at 0 K, but mechanisms can change at higher temperatures. For example, diffusion in Si at processing temperatures appears to be governed by collective atomic motions that do not operate at lower temperatures.[26] Furthermore, DFT calculations typically ignore entropic effects, some of which can change pre-exponential factors by many orders of magnitude.[27] Finally, DFT is a ground-state theory and predicts bulk bandgaps very poorly. By implication, deep electronic levels associated with point defects incur similar errors.[28,29]

For pure Si dimer dissociation, Pelaz *et al.*[24] reported a barrier of 1.6 eV by fitting experimental SIMS data for a delta-doped superlattice. From DFT calculations, Kim *et al.*[30] obtained dissociation energies of 1.30 eV using the local density approximation (LDA) and 1.36 eV using the generalized gradient approximation (GGA). Further, Bongiorno *et al.*[31] employed tight binding molecular dynamics simulations to calculate a barrier of 1.41 eV. From these results, ML estimation gives $E_2 = 1.41 \pm 0.03$ eV for Si dimer dissociation into two free Si interstitials.

For pure B dimer $(B_i - B_s)$ dissociation, Zhu *et al.*[32] used DFT (LDA) to obtain a binding energy of 1.8 ± 0.1 eV. Tight binding molecular dynamics studies by Luo *et al.*[33] yielded a corresponding number of 1.6 eV. Based on these results, ML estimation yields $E_{2,B} = 1.70 \pm 0.07$ eV for boron dimer dissociation into B_i and B_s.

For large clusters of pure Si, we rely on the literature for {311}-defect dissociation, even though such defects are larger than the five-atom clusters allowed by FLOOPS and probably have larger dissociation energies.[24,25] We justify this choice as follows. During spike annealing in the presence of many kinds of clusters whose dissociation energies vary with size, the instantaneous free interstitial concentration at any moment during heating is controlled primarily by the number of clusters dissociating at that time. When dissociation energies increase monotonically with cluster size (or at least approximately so), dissociating one atom from a cluster sets off a rapid cascade of further dissociation events that quickly liberates all the interstitials in the cluster. Equating the five-atom dissociation energy to that for very large defects merely telescopes that cascade

Table II. Activation energies for cluster dissociation.[a]

Composition	Cluster size	Species liberated	Symbol	Activation Energy (eV)	Reference	Method[b]
Pure B	2	B	$E_{2,B}$	1.70 ± 0.07	This work	ML
Pure Si	2	Si	E_2	1.41 ± 0.03	This work	ML
	3	Si	E_3	2.2	This work	Linear interpolation
	4	Si	E_4	3.0	This work	Linear interpolation
	5	Si	E_{large}	3.7 ± 0.1	This work	ML
Mixed B-Si	2[c]	B	$E_{2,mix \rightarrow B} = E_{ko}$	0.50	This work	From dopant activation[d]
	2[e]	Si	$E_{2,mix \rightarrow Si} = E_{dis}$	0.59 ± 0.06	4	ML
	3	B	$E_{3,mix}$	2.2	This work	Assumed $=E_3$
	3	Si	$E_{3,mix}$	2.2	This work	Assumed $=E_3$
	4	B	$E_{4,mix}$	3.0	This work	Assumed $=E_4$
	4	Si	$E_{4,mix}$	3.0	This work	Assumed $=E_4$
	5	B	$E_{large,mix}$	3.5	17	DFT
	5	Si	$E_{large,mix}$	3.5	17	DFT

[a] All pre-exponential factors are assumed equal to 6×10^{12} s^{-1}.[4]
[b] ML = Maximum likelihood estimation.
[c] This represents the kick-out reaction $(B_s - Si_i)^+ \rightarrow B_i^+ + Si_s$.
[d] ML method in Ref. 4 yielded 1.05 eV. The value in the table is calculated from published data for dopant activation (equivalent to solid solubility) as discussed in text.
[e] This represents the dissociation reaction $(B_s - Si_i)^+ \rightarrow B_s^- + Si_i^{+2}$.

slightly. However, equating the five-atom dissociation energy to its actual value in nature would leave no clusters with dissociation energies equal to that for large clusters. All clusters would therefore dissolve prematurely in the simulations, leading to a flood of free interstitials that would accelerate TED unnaturally.

For {311} defects, we are aware of two experimental reports of dissociation energies. Using inverse modeling of cluster evaporation applied to SIMS profiles, Rafferty *et al.*[34] reported a value of 3.57 eV. Stolk *et al.*[23] examined decay kinetics of individual {311} defects by transmission electron microscopy, obtaining 3.8 ± 0.2 eV. Maximum Likelihood estimation then gives $E_{large} = 3.7 \pm 0.1$ eV for dissociation of {311} defects and, for our case, large (five-atom) clusters of pure Si.

For large B-Si mixed clusters, we again rely on the literature for very large clusters for reasons analogous to those given for pure Si clusters. We are aware of only one relevant report by Liu *et al.*[15] who used DFT (GGA) for to obtain a per-atom binding energy of 3.5 eV for $B_{12}Si_7$. The "per-atom" energy is averaged over all atoms including both B and Si, so the calculation cannot be used to differentiate dissociation that liberates free Si *vs.* free B. With only one report available, ML estimation becomes trivial and yields $E_{large,mix} = 3.5$ eV for large (five-atom) mixed B-Si clusters, and therefore for five-atom mixed clusters of any stoichiometry from BSi_4 to B_4Si releasing either B or Si free interstitials.

The remaining dissociation energies for trimers and tetramers of pure Si and mixed B-Si were estimated using simple assumptions made in the spirit of the ML estimation. For pure Si clusters, the trimer and tetramer energies E_3 and E_4 were calculated by simple linear interpolation of the ML numbers for dimers and pentamers. The energies $E_{3,mix}$ and $E_{4,mix}$ for mixed B-Si clusters of any stoichiometry, releasing either B or Si free interstitials, were then set equal to these interpolations.

Kick-out and boron solubility.—In order to obtain a satisfactory match to experimental dopant activation data, we found it necessary to relax ML constraints on one parameter: the activation energy for kick-out $E_{2,mix \rightarrow B}$. The ML result has previously been calculated to be 1.05 eV,[4] but we found it necessary to reduce this value to 0.5 eV. Other parameters in principle could have been chosen for adjustment, but we chose this one based on its origin in DFT reports from only one laboratory[35,36] (compared with a larger suite of reports for closely related parameters) and the curiously large values of these reports compared with other activation energies associated with this complex.

We obtained the value of 0.50 eV in the following way. Experimental investigations of the spreading resistance profile (SRP) showed that the degree of boron activation after a spike anneal roughly equals the solid solubility at the peak temperature.[37,38] This solubility lies near $[B_s^-] = 10^{20}$ cm^{-3}.[2] Substitutional boron is formed from the complex $(B_s - Si_i)^+$. The time constants for exchange of B_s with the surroundings through this complex are much faster than those of profile spreading overall, so a quasi-steady state mass balance can be written for $(B_s - Si_i)^+$

$$
\begin{aligned}
0 &= r_{ki} - r_{2,mix \rightarrow B} + r_{assoc} - r_{2,mix \rightarrow Si} \\
&= k_{ki}[Bi_i^+] - k_{2,mix \rightarrow B}[(B_s - Si_i)^+] + k_{assoc}[B_s^-][Si_i^{+2}] \\
&\quad - k_{2,mix \rightarrow Si}[(B_s^- Si_i)^+] \qquad [7]
\end{aligned}
$$

Although in principle, Eq. 7 includes contributions from interstitial exchange with trimers, the rate expressions in the tables together with likely estimates for trimer concentrations (confirmed by simulations) indicate that exchange of interstitials between $(B_s - Si_i)^+$ and various mixed trimers is negligible. Our simulations lead to $[(B_s - Si_i)^+] \approx 1 \times 10^{13}$ cm^{-3}, with $[Si_i^{+2}] \approx 5 \times 10^{15}$ cm^{-3} and $[B_i^+] \approx 1 \times 10^{13}$ cm^{-3} near the surface at 1000°C during ramp up. (Calculations done at different positions and temperatures gave similar results.) With ML parameters for E_{ki}, $E_{2,mix \rightarrow Si}$, and

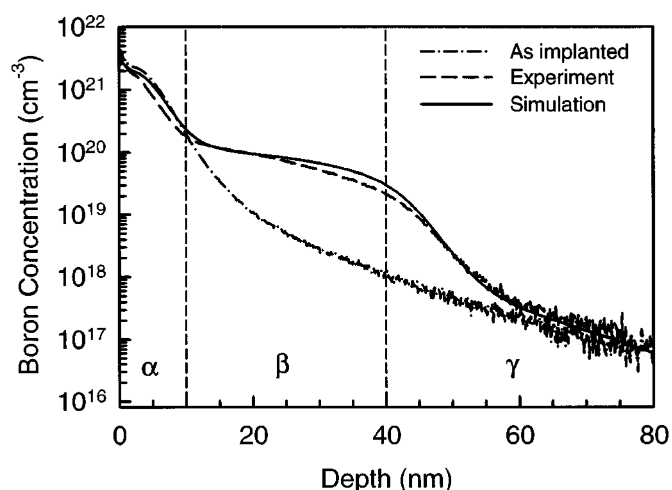

Figure 2. Experimental SIMS profiles for boron implanted at 0.60 keV, 2×10^{15} ions/cm^2, and 0° tilt. As-implanted profile is shown together with the results after a spike anneal to 1050°C at a heating rate of 150°C/s. For descriptive purposes, the profile is divided into α,β, and γ regions. Also shown is a simulated profile for "best" parameter values selected mainly by the ML estimation and listed in Tables I and II.

E_{assoc} drawn from Table I and II, $E_{2,mix\rightarrow B}$ calculated from Eq. 7 becomes 0.5 eV. Figure 2 shows an example of the excellent agreement this choice for $E_{2,mix\rightarrow B}$ gives with experimental SIMS profiles.

Results

Sensitivity coefficients S were computed according to Eq. 2 using the following vector $\boldsymbol{\theta}$ of activation energies

$$\boldsymbol{\theta} = \begin{bmatrix} E_{diff,B_i^+} \\ E_{diff,Si_i^{+2}} \\ E_{ki} \\ E_{2,mix\rightarrow B} \\ E_{2,mix\rightarrow Si} \\ E_{2,B} \\ E_2 \\ E_3 \\ E_4 \\ E_{large} \\ E_{large,mix} \end{bmatrix} \quad [8]$$

Table I and II show in compact form the definitions of each vector element. We used FLOOPS to calculate boron profiles according to the TED model described earlier. One dependent variable y was the junction depth X_j, defined as the depth at which the boron concentration decreases to 10^{18} cm^{-3}. Sensitivity coefficients connected with junction depth are denoted by $S_{jxn,j}$. A second dependent variable was the fraction of electrically active boron (in substitutional sites). Sensitivity coefficients connected with boron activation are denoted by $S_{B\text{-}activ,j}$.

We calculated each value of S_{ij} three times, using deviations $\Delta\theta_j$ of 5, 10, and 20% in either direction from the independent variable's nominal value given in Table I and II. Although using a variation as large as 20% degrades the accuracy of the finite difference approximation to the exact value of S_{ij} shown in Eq. 1, using this large variation in concert with a smaller one offers a simple way to detect threshold effects and other nonlinearities in the model that can lead to important physical insights. Little variation in S_{ij} with the magnitude or direction of the deviation shows that the model is in a well-behaved, approximately linear regime with respect to the parameter j. When the variations in S_{ij} are strong, however, the model is in or very near to a significantly nonlinear region where the importance of parameter j is changing rapidly. In a kinetic model, such nonlinearity often means that a mechanism is changing or some new elementary step is becoming important.

Table III shows the calculated sensitivity coefficients. A negative value for S_{ij} indicates that an increase in the independent variable j led to a decrease in the dependent variable i. The highest and lowest sensitivities differ by roughly three orders of magnitude for $S_{jxn,j}$ and four orders of magnitude for $S_{B\text{-}activ,j}$, showing that the various activation energies exert effects of vastly different importance on the diffusion profiles. Note that the two groups of sensitivity coefficients for the two corresponding dependent variables of junction depth and dopant activation should not be compared directly because the coefficient groups have different units. Valid comparisons are possible only for coefficients within a given group. The absolute magnitudes of the coefficients indicate how much the dependent variable changes in response to a unit change in the given parameter, while the signs describe the directionalities of these changes. Although this absolute magnitude may be useful for some purposes, this paper concerns itself primarily with the relative magnitudes of the coefficient groups for each dependent variable. Within each group, activation energies having the largest coefficients on a relative scale exert the strongest effects on the corresponding dependent variable.

Table III. Sensitivity coefficients S_{ij}.

Independent variable	$S_{jxn,j}$ (nm/eV)			$S_{B\text{-}activ,j}$ (eV^{-1})		
	5%[a]	10%[a]	20%[a]	5%[a]	10%[a]	20%[a]
E_{diff,B_i}	-2.9×10^2	-2.7×10^2	-2.5×10^2	1.3	1.1	1.2
E_{diff,Si_i}	<2[b]	<2	<2	$<1 \times 10^{-2}$[c]	$<1 \times 10^{-2}$	$<1 \times 10^{-2}$
E_{ki}	2.8×10^2	2.6×10^2	2.4×10^2	−1.2	−1.1	−1.1
$E_{2,mix\rightarrow B}$	-2.7×10^2	-2.6×10^2	-2.4×10^2	1.1	1.0	9.7×10^{-1}
$E_{2,mix\rightarrow Si}$	2.8×10^2	2.6×10^2	2.5×10^2	−1.2	−1.1	−1.2
$E_{2,B}$	-6.1×10^1	-5.9×10^1	-5.8×10^1	1×10^{-2}	1×10^{-2}	3×10^{-2}
E_2	<2	<2	<2	$<1 \times 10^{-2}$	$<1 \times 10^{-2}$	$<1 \times 10^{-2}$
E_3	-6.4×10^1	-6.1×10^1	-6.0×10^1	3×10^{-2}	3×10^{-2}	2×10^{-2}
E_4	-6.4×10^1	-6.2×10^1	-6.2×10^1	$<1 \times 10^{-2}$	$<1 \times 10^{-2}$	-2×10^{-2}
E_{large}	−3	−8	-2.5×10^1	$<1 \times 10^{-2}$	$<1 \times 10^{-2}$	2×10^{-2}
$E_{large,mix}$	−4	−3	-2.0×10^1	-6.4×10^{-1}	-6.6×10^{-1}	-3.7×10^{-1}

[a] Percentage signifies the fractional variation in the independent variable $\boldsymbol{\theta}_j$ used to obtain S_{ij}.
[b] The symbol <2 signifies that the magnitude is less than 2.
[c] The symbol $<1 \times 10^{-2}$ signifies that the magnitude is less than 1×10^{-2}.

Discussion

Effects on junction depth.—The junction depth responds strongly to variations in $E_{\mathrm{diff,B_i^+}}$ and to three activation energies associated with forming and dissociating the mixed B-Si dimer complex $(B_s - Si_i)^+$. These latter energies include $E_{2,\mathrm{mix}\to B}$ (kick-out to liberate B_i), $E_{2,\mathrm{mix}\to \mathrm{Si}}$ (dissociation to create B_s and Si_i), and E_{ki} (kick-in). The importance of the dimer results from its activity as a reservoir with which B_i exchanges during a temperature spike. Profile spreading takes place primarily in response to motion of B_i (as opposed to diffusion of $(B_s - Si_i)^+$ itself[4]). The $(B_s - Si_i)$ complex forms and then quickly redissociates by liberating either B_i or Si_i. Liberation of B_i frees the boron atom again. Liberation of Si_i immobilizes the boron as B_s, however, until another Si_i comes along to reform the $(B_s - Si_i)$ complex and provide a mechanism for releasing B_i. The branching ratio between the dissociation reactions forming B_i and Si_i determines what fraction of time boron is free to move, and $E_{\mathrm{diff,B_i^+}}$ determines how far the freed boron interstitial can move before it becomes captured by $(B_s - Si_i)^+$ again.

There is another primary kind of reservoir to consider that could appreciably hamper motion of B_i: nondissociated clusters. The importance of the nondissociated clusters scales with their concentration and dissociation energies. The larger sensitivity coefficients for the activation energies associated with forming and dissociating the mixed B-Si dimer complex suggests that this mechanism plays a more important role than nondissociated clusters in determining the diffusion distance of interstitial boron.

Other parameters having appreciable sensitivity coefficients for junction depth include dissociation energies for boron-containing clusters of small to medium size. Sensitivity coefficients for a given cluster size increase with the ability to liberate mobile B_i over the course of the annealing cycle. Small and medium-sized clusters have lower dissociation energies than large clusters and dissociate nearly completely. As long as large clusters do not dissociate, the exact value of their dissociation energy has little bearing on the release of B_i, so the corresponding sensitivity coefficients remain small.

The ability of large clusters to liberate B_i does increase, however, under conditions in which they become active. Thus, the sensitivity coefficients for the largest clusters should increase as their dissociation energy is varied downward, or equivalently, the maximum spike temperature increases. Evidence for this effect can be seen in Table III by comparing sensitivity coefficients for $\Delta\theta$ of 10 and 20%. The 10% variations lead to little sensitivity at the largest cluster sizes because those clusters have dissociation energies too large to become active at 1050°C. The 20% variations in $\Delta\theta$ lead to corresponding sensitivity coefficients that are a factor of ten larger for $E_{\mathrm{large,mix}}$. These clusters have been assigned a dissociation energy of 3.5 eV, which lies just on the cusp of becoming active at the maximum temperature of 1050°C. The negative variations in $\Delta\theta$ lead to greatly increased dissociation fractions at 1050°C.

The exact values of the sensitivity coefficients are influenced by other factors as well. For example, sensitivities should vary in proportion to the number of clusters available at a given size. The number of clusters depends upon the initial size distribution obtained after annealing; very large clusters with small concentrations would have very small sensitivity coefficients. Such numerical effects are amplified by stoichiometry; for example, size four clusters liberate twice as many boron atoms on average as size two clusters. The present simulations may capture size distribution effects imperfectly due to the small maximum cluster size employed (five). The maximum dissociation energy is probably not reached until sizes much larger than five, as suggested by published work.[24,25]

Interestingly, junction depth is insensitive to parameters connected with pure Si species, notably $E_{\mathrm{diff,Si_i^{+2}}}$, E_2, and E_{large}. Although the sensitivity coefficients E_3 and E_4 are more significant, these parameters describe both pure Si and mixed B-Si clusters-the effect probably arises from the mixed clusters. Regarding $E_{\mathrm{diff,Si_i^{+2}}}$, the simulations show that the concentration profiles for Si_i tend to be relatively flat near the junction region because there is no mechanism equivalent to kick-in for boron to help sequester Si_i. Exchange of Si_i with Si in the lattice simply yields more Si_i. The rapid motion tends to flatten the profiles and make the exact value of the interstitial diffusivity immaterial. From a simulation standpoint, the unimportance of $E_{\mathrm{diff,Si_i^{+2}}}$ is helpful because of the great uncertainty[39] in its value. Regarding dissociation energies for pure Si clusters, these clusters exert an influence on junction depth only indirectly by releasing Si_i, which accelerates the kick-out reaction of B_s. Apparently the details of this rate of release do not affect junction depth very much. This result suggests that nonmonotonic variations in Si cluster dissociation energy with size, which has been observed in some experiments,[25] may not play a significant role. Presumably analogous variations for mixed B-Si clusters would have larger effects, however, because of their capacity to release B_i directly.

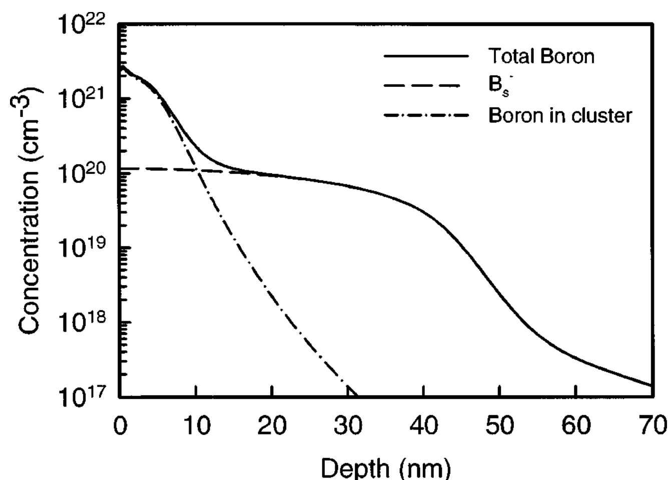

Figure 3. Simulation profiles of various species of boron. Electrically inactive clusters dominate in the α region, but the β and γ regions have approximately 100% dopant activation.

Effects on dopant activation.—Dopant activation integrated over the entire profile typically ranged between 20 and 40% in these simulations. As shown in Fig. 3, activation in the β and γ regions was typically near 100%, falling to lower levels only in the α region where boron concentrations rose above 10^{20} cm^{-3} and B-containing clusters were prevalent. Thus, variations in overall activation tended to result almost entirely from changes in the transition zone between the α and β regions.

Dopant activation responds most strongly to variations in the same parameters as junction depth. This result can be rationalized by recognizing that boron becomes electrically active only by entering lattice sites via the $(B_s - Si_i)^+$ complex and then dissociating to liberate Si^{+2}. The dominance of $E_{\mathrm{diff,B_i^+}}$, $E_{2,\mathrm{mix}\to B}$, $E_{2,\mathrm{mix}\to \mathrm{Si}}$, and E_{ki} over the other parameters (except for $E_{\mathrm{large,mix}}$) is much more pronounced than it is for junction depth, probably because there is no competing pathway for forming substitutional boron.

Interestingly, among all the cluster dissociation energies, only $E_{\mathrm{large,mix}}$ shows a sensitivity coefficient even remotely comparable to the four just discussed. The computed sensitivity coefficient varies by nearly a factor of two as $\Delta\theta$ increases from 10 to 20%. This cluster begins to become a major supplier of interstitial boron at 1050°C, especially in the α region (away from the junction). Thus the simulations behave rather sensitively and nonlinearly to variations in the corresponding dissociation energy. Clearly, accurate models of dopant activation need to pay special attention to estimating this parameter correctly, as has been suggested in the literature.[23]

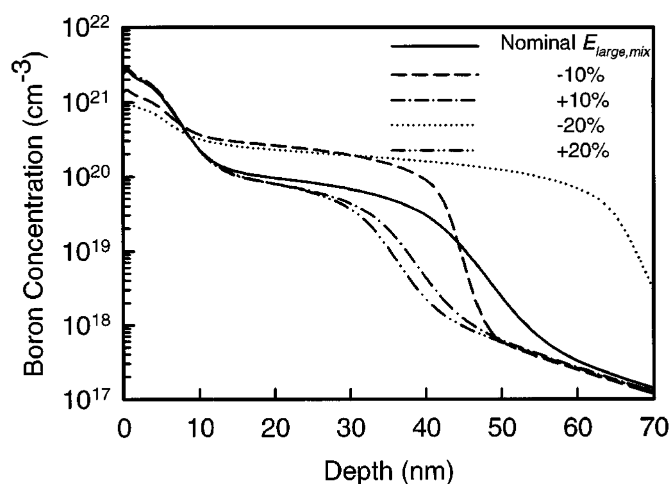

Figure 4. Comparison of simulated profiles from variations in $E_{\mathrm{large,mix}}$.

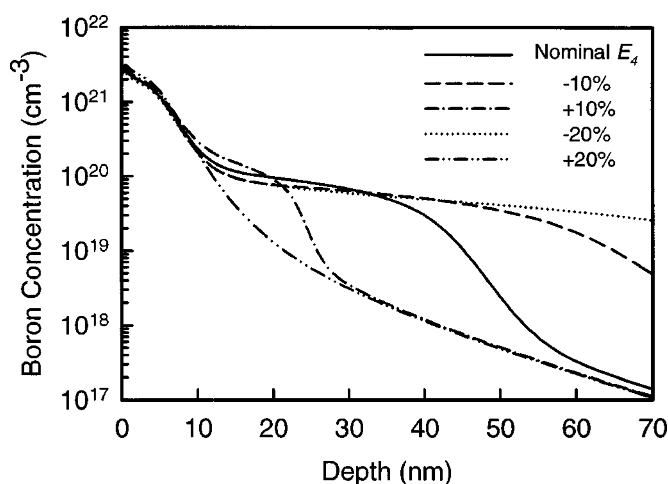

Figure 6. Comparison of simulated profiles from variations in E_4.

Effects on profile shape.—The various activation energies influence other aspects of profile shape in ways that are not adequately captured by a single metric like junction depth. Examples include the transition depth between the α, β, and γ regions in Fig. 2 as well as the dopant concentration in these regions. Although metrics can be defined for these profile features, it seems more instructive to show the effects through the profiles themselves.

The effects of $E_{\mathrm{large,mix}}$ were especially pronounced, as shown in Fig. 4. As expected, the largest effects appeared for negative variations because the cluster liberates more boron, and the effects extended throughout the profile including the α region.

Other parameters influenced only the β and γ regions deeper in the bulk. Parameters that behaved this way include $E_{\mathrm{diff,Bi^+}}$, E_{ki}, $E_{2,\mathrm{mix}\rightarrow\mathrm{B}}$, $E_{2,\mathrm{mix}\rightarrow\mathrm{B}}$, $E_{2,\mathrm{B}}$, and E_3. Typical behavior appears in Fig. 5 for $E_{\mathrm{diff,Bi^+}}$. For E_{ki}, the magnitude in horizontal separation in the γ region and vertical separation in the β region resembles those in Fig. 5. For $E_{2,\mathrm{mix}\rightarrow\mathrm{B}}$ and $E_{2,\mathrm{mix}\rightarrow\mathrm{B}}$, the effects were somewhat more pronounced. The effects were more subdued for $E_{2,\mathrm{B}}$ and E_3, especially in the β region. In all these cases, variation in the activation energy caused smooth changes in the position of the break point between the β and γ regions, together with smooth changes in the average dopant concentration in the β region. Raising and lowering the profile in the β region, also has the consequence of moving the break point between the α and β regions somewhat.

For E_4, the situation changed as shown in Fig. 6. The transition between the β and γ regions varied so strongly that for a 20% increase in E_4, the shoulder disappeared entirely. Small variations in the α region appear as well. By contrast, the effects of E_{large} are small in most simulations of Fig. 7 and appear only in the γ region, except for the case of a 20% decrease, where the significant variations show up in both the β and γ regions. As mentioned earlier, very few large clusters of pure Si dissociated during a spike to 1050°C unless E_{large} was adjusted significantly downward.

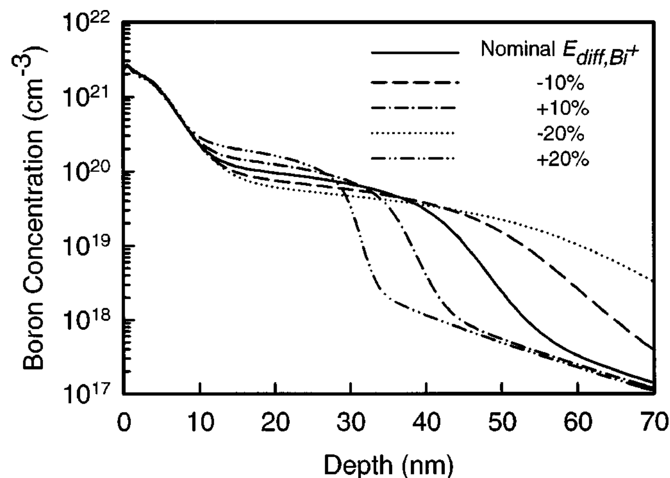

Figure 5. Comparison of simulated profiles from variations in $E_{\mathrm{diff,B_i^+}}$. Similar profiles are observed for the variations in E_{ki}, $E_{2,\mathrm{mix}\rightarrow\mathrm{B}}$, $E_{2,\mathrm{mix}\rightarrow\mathrm{B}}$, $E_{2,\mathrm{B}}$ and E_3.

Conclusion

Sensitivity analysis coupled with simple local kinetic approximations have offered insights into the dominant mechanisms that govern junction depth and dopant activation in TED. For example, the larger sensitivity coefficients for the activation energies associated with forming and dissociating the mixed B-Si dimer complex suggests that this mechanism plays a more important role than nondissociated clusters in determining the degree of profile spreading. The curious behavior of the sensitivity coefficient for dissociation energy of the largest mixed clusters suggests that the dissociation energy and initial size distribution may play a key role in determining the final degree of dopant activation. Clearly these mechanisms deserve

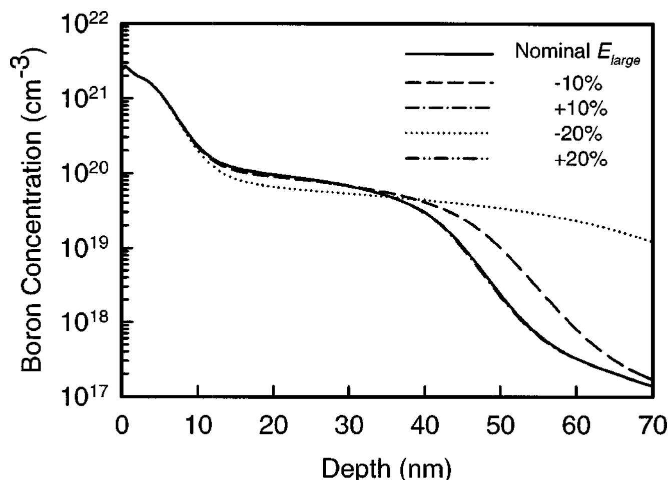

Figure 7. Comparison of simulated profiles from variations in E_{large}.

the most attention from future experiments and computations designed to obtain elementary rate parameters of interest in post-implant annealing.

Acknowledgments

This work was supported by NSF (CTS 98-06329) and by Sematech.

References

1. See, for example, the *International Technology Roadmap for Semiconductors*, http://public.itrs.net/Files/2001ITRS/Home.htm
2. *Rapid Thermal Processing: Science and Technology*, R. B. Fair, Editor, Academic Press, Boston (1993).
3. See, for example, *Mater. Res. Soc. Symp. Proc.,* **568** (1999); *Mater. Res. Soc. Symp. Proc.,* **610** (2000).
4. M. Y. L. Jung, R. Gunawan, R. D. Braatz, and E. G. Seebauer, *AIChE J.*, Submitted.
5. W. Lerch, M. Gluck, N. A. Stolwijk, H. Walk, M. Schafer, S. D. Marcus, D. F. Downey, and J. W. Chow, *J. Electrochem. Soc.,* **146**, 2670 (1999).
6. R. Tomovic and M. Vukobratovic, *General Sensitivity Theory*, Elsevier, New York (1972).
7. P. M. Frank, *Introduction to System Sensitivity Theory*, Academic Press, New York, (1978).
8. A. Varma, M. Morbidelli, and H. Wu, *Parametric Sensitivity in Chemical Systems*, Cambridge University Press, New York (1999).
9. See M. Law, http://www.swamp.tec.ufl.edu/
10. H.-H. Vuong, C. S. Rafferty, S. A. Eshraghi, J. Ning, J. R. McMacken, S. Chaudhry, J. McKinley, and F. A. Stevie, *J. Vac. Sci. Technol. B,* **18**, 428 (2000).
11. M. Y. L. Jung, R. Gunawan, R. D. Braatz, and E. G. Seebauer, *J. Appl. Phys.*, Submitted.
12. M. J. Caturla, M. Foad, and T. D. de la Rubia, in Proceedings of the International Conference on Ion Implantation Technology, Vol. 2, IEEE, p. 1018 (1999).
13. H. Kobayashi, I. Nomachi, S. Kusanagi, and F. Nishiyama, *Mater. Res. Soc. Symp. Proc.,* **669**, J 5.3 (2001).
14. D. F. Downey, S. W. Falk, A. F. Bertuch, and S. D. Marcus, *J. Electron. Mater.,* **28**, 1340 (1999).
15. L. H. Zhang, K. S. Jones, P. H. Chi, and D. S. Simons, *Appl. Phys. Lett.,* **67**, 2025 (1995).
16. M. J. Caturla, M. D. Johnson, and T. D. de la Rubia, *Appl. Phys. Lett.,* **72**, 2736 (1998).
17. X.-Y. Liu, W. Windl, M. P. Masquelier, *Appl. Phys. Lett.,* **77**, 2018 (2000).
18. See, for example, K. J. Laidler, *Chemical Kinetics*, Wiley & Sons, New York (1980).
19. M. M. Bunea, P. Fastenko, and S. T. Dunham, *Mater. Res. Soc. Symp. Proc.,* **568**, 135 (1999).
20. C. S. Rafferty, G. H. Gilmer, M. Jaraiz, D. Eaglesham, and H.-J. Gossman, *Appl. Phys. Lett.,* **68**, 2395 (1996).
21. S. Chkravarthi and S. T. Dunham, *J. Appl. Phys.,* **89**, 3650 (2001).
22. B. J. Mulvaney and W. B. Richardson, *J. Appl. Phys.,* **67**, 3197 (1990).
23. P. A. Stolk, H.-J. Gossmann, D. J. Eaglesham, D. C. Jacobson, C. S. Rafferty, G. H. Gilmer, M. Jaraíz, J. M. Poate, H. S. Luftman, and T. E. Haynes, *J. Appl. Phys.,* **81**, 6031 (1997).
24. L. Pelaz, M. Jaraíz, G. H. Gilmer, H.-J. Gossmann, C. S. Rafferty, D. J. Eaglesham, and J. M. Poate, *Appl. Phys. Lett.,* **70**, 2285 (1997).
25. 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. Jaraíz, *Phys. Rev. Lett.,* **82**, 4460 (1999).
26. J. A. Van Vechten, *Phys. Rev. B,* **38**, 9913 (1988).
27. J. A. Van Vechten and C. D. Thurmond, *Phys. Rev. B,* **14**, 3539 (1976).
28. W. Windl, M. M. Bunea, R. Stumpf, S. T. Dunham, and M. P. Masquelier, *Phys. Rev. Lett.,* **83**, 4345 (1999).
29. W. Windl, O. F. Sankey, and J. Menendez, *Phys. Rev. B,* **57**, 2431 (1998).
30. J. Kim, F. Kirchhoff, W. G. Aulbur, J. W. Wilkins, F. S. Khan, and G. Kresse, *Phys. Rev. Lett.,* **83**, 1990 (1999).
31. A. Bongiorno, L. Colombo, F. Cargnoni, C. Gatti, and M. Rosati, *Europhys. Lett.,* **50**, 608 (2000).
32. J. Zhu, T. D. de la Rubia, L. H. Yang, C. Mailhiot, and G. H. Gilmer, *Phys. Rev. B,* **54**, 4741 (1996).
33. W. Luo, P. B. Rasband, P. Clancy, and B. W. Roberts, *J. Appl. Phys.,* **84**, 2476 (1998).
34. C. S. Rafferty, G. H. Gilmer, M. Jaraíz, D. Eaglesham, and H.-J. Gossman, *Appl. Phys. Lett.,* **68**, 2395 (1996).
35. J. Zhu, T. Diaz de la Rubia, L. H. Yana, C. Mailhiot, and G. H. Gilmer, *Phys. Rev. B,* **54**, 4741 (1996).
36. J. Zhu, *Comput. Mater. Sci.,* **12**, 309 (1998).
37. E. Schroer, V. Privitera, F. Priolo, E. Napolitani, A. Carnera, and S. Moffatt, *Mater. Sci. Eng., B,* **71**, 219 (2000).
38. E. J. H. Collart, A. J. Murrell, M. A. Foad, J. A. van den Berg, S. Zhang, D. Armour, R. D. Goldberg, T.-S. Wang, A. G. Cullis, T. Clarysse, and W. Vandervorst, *J. Vac. Sci. Technol. B,* **18**, 435 (2000).
39. D. Eaglesham, *Phys. World,* **8**, 41 (1995).