Maximum *A Posteriori* Estimation of Transient Enhanced Diffusion Energetics

R. Gunawan, M. Y. L. Jung, E. G. Seebauer, and R. D. Braatz
Dept. of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Transient enhanced diffusion (TED) of boron limits the formation of ultrashallow junctions needed in next-generation microelectronic devices. A comprehensive TED model needs many parameters governing the physical and chemical processes. Prior estimates of the most likely values for the parameters as well as their accuracies are determined from maximum likelihood estimation applied to estimates from focused individual experiments and density functional theory calculations. Here a systematic approach to model-parameter identification using maximum a posteriori estimation is employed combining the maximum likelihood parameter estimates and their uncertainties in conjunction with after-anneal boron SIMS profiles to obtain accurate TED energetics. Guidance on future experimental and ab initio efforts are given based on the agreement (and disagreement) between the prior and posterior distributions.

Introduction

Technological advances in packing an ever-increasing number of transistors into a single chip drive the evolution in the microelectronics industry. Because future Si-based CMOS requires formation of junction depths between 13 and 25 nm in the source and drain extension regions by the year 2005 according to the 2001 International Technology Roadmap for Semiconductors, problems accompanying ultrashallow junction processing have become increasingly important. The current technology for forming ultrashallow junctions relies on using ion implantation to introduce dopant into the wafer. However, ion implantation causes lattice damage, producing junctions with high sheet resistance, which motivates postimplant annealing to remove the implant damage. Transient enhanced diffusion (TED) occurs during annealing of ion-implanted Si in which dopant, especially boron, diffuses exceptionally fast, resulting in an undesirable increase of junction depth. TED is often attributed rather generically to the supersaturation of Si interstitials produced during ion implantation of dopant and during thermal annealing. However, obtaining a detailed kinetic understanding of how the Si interstitials mediate dopant diffusion is more challenging. Furthermore, as thermal annealing advances toward rapid thermal annealing (RTA), it becomes crucial to implement more precise kinetic expressions to capture the transient behavior. For this reason, modeling of TED has generated a great deal of interest (see, for example, Agarwal et al. (2000)).

Here, a comprehensive TED model is developed combining elementary physicochemical processes (such as diffusion of interstitials and boron activation reaction) and Poisson's equation. The model includes continuity equations describing Fickian diffusion and electric drift motion, and formation and annihilation rates due to chemical reactions (such as dopant activation, dopant-defect clustering). Poisson's equation describes the electric field due to spatial imbalance of the charge density. The TED model consists of 20 coupled stiff nonlinear partial and ordinary differential equations.

An accurate TED model requires knowledge of many kinetic parameters describing various physics. These parameters include diffusivities and kinetic rate constants whose thermal behavior is typically assumed to follow the Arrhenius law. The values of the Arrhenius parameters (that is, the activation energies) reported in the literature, which have been obtained through either carefully designed focused experiments or from density functional theory calculations, can vary by an order-of-magnitude. To resolve the disagreement in the literature, maximum likelihood (ML) estimation (Beck and Arnold, 1977) was employed using published values of activation energies giving rationally defensible estimates for most of the parameters (Gunawan et al., 2002a). ML estimation gave the most likely parameter values as well as their accu-

Correspondence concerning this article should be addressed to R. D. Braatz.

racy estimates. The parameters and their corresponding accuracies provide prior distributions for further parameter estimation studies, for example, using Bayesian estimation (Beck and Arnold, 1977).

Maximum *a posteriori* (MAP) estimation (Beck and Arnold, 1977) has been extensively applied in many field of studies, for example, in imaging (Xiao et al., 2002), signal processing (Zarnich et al., 2001), pattern recognition (Chang and Park, 2001), and spectral estimation (Lamberg et al., 2001). MAP estimation builds upon Bayes' theorem, which describes *a posteriori* distribution of the parameter estimates as a function of *a priori* distribution function and experimental observations. The general idea behind MAP estimation is to optimally combine prior information of the most likely parameter values with additional experimental data to obtain better estimates.

In this work, we employ a systematic approach to model-parameter identification using MAP estimation that combines prior parameter information from published parameter values and experimental data from after-anneal boron secondary ion mass spectroscopy (SIMS) profiles to give accurate estimates of TED parameters. MAP estimation is formulated as a minimization problem that is numerically solved using the sequential quadratic programming package CFSQP (Lawrence et al., 1997). The TED model is implemented using the process simulator FLOOPS 2000 (Law and Cea, 1998). Comparison of the *a posteriori* and *a priori* distributions provides information for evaluating the quality of the prior parameter estimates obtained by focused experiments and density functional theory calculations. Guidance on future experimental and *ab initio* efforts are given based on the agreement (and disagreement) between the prior and posterior distributions.

Transient Enhanced Diffusion Model

Transient enhanced diffusion arises from coupled physical processes (such as diffusion and electric drift motion) and chemical reactions (such as boron activation and cluster formation and dissolution). The model consists of continuity equations describing Fickian diffusion and electric drift motion for mobile species, and formation and annihilation rates due to chemical reactions. The mass balances/continuity equations are

$$\frac{\partial N_i}{\partial t} = -\frac{\partial J_i}{\partial x} + G_i \tag{1}$$

where N_i denotes the concentration, J_i denotes the flux, and G_i denotes the net generation rate of species i. The flux J_i describes the motion of species due to Fickian diffusion and electrical drift (see Appendix for details). Here, the mobile species consist of boron and silicon interstitials (B_i and S_i), whereas the immobile species include substitutional boron (B_s), boron–silicon complex (B_s–Si_i), and the clusters (pure boron, pure silicon, and mixed boron–silicon clusters). The generation rate G_i includes chemical reactions associated with the boron activation and dopant-defect clustering (see Appendix for further details).

The spatial dependence of species concentrations and species charge distributions create an imbalance in the spatial charge density, which induces an electric field according to Poisson's equation

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{Q(x)}{k_d} \tag{2}$$

where ψ denotes the potential and k_d denotes the dielectric constant. The charge density $Q(x)$ describes the net charge concentration according to

$$Q(x) = p - n + \sum_i \gamma_i N_i \tag{3}$$

where p and n denote the hole and electron concentrations, respectively, and γ_i denotes the net charge of species i (see Appendix for definition). The carrier concentrations (that is, p and n) are assumed to be in thermal equilibrium at all times.

The proposed TED model is the most comprehensive to date. Other models often lump the electric field effects and/or the dependence on the supersaturation of Si_i in an enhancement scaling factor of the dopant diffusivity (Law et al., 1991; Lerch et al., 1999), while some models completely ignore the electric drift motion of dopants (Bennett and Price, 1994; Cowern et al., 1990). In addition, the treatment of clusters in existing TED models typically use a simplified approach where the dynamic of clusters was described using a single fictitious cluster concentration variable, ignoring the size-dependent cluster dissociation energies (Bennett and Price, 1994; Law et al., 1991). In addition, the clustering of dopant and defects is completely ignored in many models (Cowern et al., 1990; Lerch et al., 1999).

Problem Formulation

The TED model requires a set of preexponential factors and activation energies associated with the diffusivities and kinetic rate constants for the boron activation reaction and cluster dissociation dynamics. Theoretical and experimental works of TED have paid little attention to the preexponential factors, focusing instead on the activation energies. In fact, prefactors are notoriously difficult to estimate *a priori*. Simple kinetic models for elementary reactions view prefactors as modified attempt frequencies. Solids have a distribution of vibrational frequencies, however, so that a single aggregate attempt frequency is difficult to define. One commonly used estimate is the Debye frequency, which for Si is about 6×10^{12} s^{-1}. However, it is well known from other branches of kinetics that such a simple picture is often inadequate. For example, a survey (Wang and Seebauer, 2001) of prefactors for gas desorption from semiconductor surfaces shows that, while the average value indeed lies near 10^{13} s^{-1}, only 10% of individual cases fall within an order of magnitude of this range. The story is roughly similar for hopping diffusion on semiconductors (Seebauer and Jung, 2001). Although the Si Debye frequency will be employed in the present work, clearly the use of *a priori* estimates of prefactors requires caution.

Designing experiments in which highly accurate values for the activation energies can be directly measured is challeng-

Table 1. Maximum Likelihood Estimates of Transient Enhanced Diffusion Energetics

TED Parameters	ML Est.	Std. Dev.
$E_{\mathrm{diff,B}_i^+}$ (boron diffusivity)	0.37	0.04
$E_{\mathrm{diff,Si}_i^{+2}}$ (Si diffusivity)	0.72	0.03
E_{ko} (kick-in/kick-out)	0.5	0.07
E_{ki} (kick-in/kick-out)	0.51	0.1
E_{dis} (kick-in/kick-out)	0.59	0.06
$E_{2,\mathrm{B}}$ (cluster dissociation—pure boron)	1.7	0.07
E_2 (cluster dissociation—pure Si)	1.4	0.03
E_3 (cluster dissociation—size 3)	2.2	n/a*
E_4 (cluster dissociation—size 4)	2.9	n/a*
E_{large} (cluster dissociation—large Si)	3.7	0.1
$E_{\mathrm{large,mix}}$ (cluster dissociation—large mixed)	3.5	n/a**

*Parameter estimate came from linear interpolations of ML estimates. In the MAP estimation, a large number is used in place of the standard deviation.
**ML estimation utilized only a single published value.

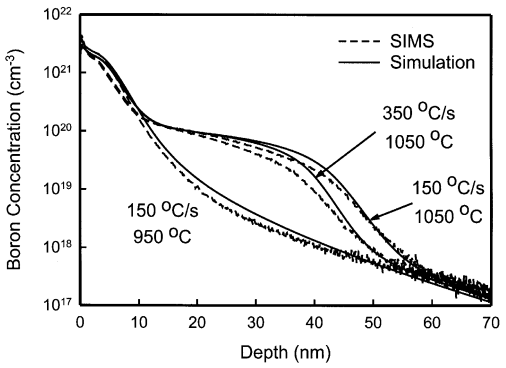

Figure 1. After-anneal SIMS and simulation profiles using ML estimates for various annealing programs.

ing. Experimental and *ab initio* density functional theory (DFT) estimates of the activation energies are scattered throughout the literature. For most of the activation energies, the published values show significant variation. For example, the reported experimental values for diffusivities for the Si self-interstitial vary by more than ten orders of magnitude (Eaglesham, 1995). The reliability of the experimental results is questionable because the mobile species including many point and extended defects are difficult to observe, as they exist at low concentration. *Ab initio* calculations based on DFT have also proven problematic for several reasons. First, most of the DFT calculations are valid only at 0 K, but the assumed mechanism may be invalid at high temperatures. For example, Si self-interstitial diffusion may be governed by collective atomic motions, which do not occur at lower temperatures (Van Vechten, 1988). Moreover, DFT calculations typically disregard the effect of entropy, which can change the preexponential factor by orders of magnitude (Van Vechten and Thurmond, 1976).

Maximum likelihood estimation provides a rationally defensible mechanism in choosing the activation energies based on literature values. Based on the literature values, ML estimation gave the most likely estimates of TED parameters (see Table 1), which provided good experimental agreement with no fitting parameters, as shown in Figure 1. In addition, the ML estimation provided estimates of the accuracy of the parameters, giving a set of *a priori* distributions for MAP estimation.

Parameter sensitivity analysis relates process behavior to parameters, which elucidates the most important dynamics governing a process. Sensitivity analysis of the junction depth and percent of boron activation (that is, fraction of boron that actively contribute to conductivity) suggests that the most important activation energies are those associated with boron diffusion and boron activation reaction (that is, $E_{\mathrm{diff,B_i}}$, E_{ko}, E_{ki}, and E_{dis}) (Gunawan et al., 2002a). In addition, the kinetic parameters that have moderate sensitivities include the dissociation parameters for pure boron clusters and intermediate-sized clusters (that is, $E_{2,\mathrm{B}}$, E_3, and E_4). A preliminary investigation using different sets of parameters shows that combinations of the most and moderately sensitive parameters are needed to accurately simulate TED (see "Results and Discussion" section). Hence, the current study will focus on these parameters.

In the present MAP estimation, the parameter estimates and the experimental observations consist of the activation energies and the postannealing SIMS profiles, respectively. The MAP estimation can be equivalently posed as a minimization problem (see Appendix for derivation)

$$\min_{\beta}\left\{(\boldsymbol{Y}-\boldsymbol{P})^T\boldsymbol{V}_{\epsilon}^{-1}(\boldsymbol{Y}-\boldsymbol{P})+(\boldsymbol{\beta}-\boldsymbol{\mu})^T\boldsymbol{V}_{\mu}^{-1}(\boldsymbol{\beta}-\boldsymbol{\mu})\right\} \quad (4)$$

where $\boldsymbol{\beta}$ denotes the parameter estimates, $\boldsymbol{Y}$ is the experimental observations, $\boldsymbol{P}$ is the model prediction, $\boldsymbol{\mu}$ denotes the *a priori* parameter estimates, and $\boldsymbol{V}_{\epsilon}$ and $\boldsymbol{V}_{\mu}$ are the variances of the experimental data and prior parameter estimates, respectively. Since each SIMS profile contains hundreds of correlated data points describing the spatial distribution of total boron concentration (compared to only seven for the *a priori* parameters), its inclusion in the likelihood function should avoid undermining the *a priori* information (such as using weighting factors). A similar problem arises in the field of speech recognition, where the acoustic and language model information (representing the observations and the *a priori* information) need balancing through incorporation of a language weight factor in the corresponding MAP objective function. The weight factor corrects the probability estimate to account for the correlation between the neighboring acoustic frames (Bahl et al., 1980). Following the same approach, reformulation of the objective function of MAP estimation S_{MAP} gives

$$\sum_{i=1}^{d}\frac{1}{w_i}\left[\boldsymbol{Y}_i-\boldsymbol{P}(\boldsymbol{X}_i,\boldsymbol{\beta})\right]^T\boldsymbol{V}_{\epsilon_i}^{-1}\left[\boldsymbol{Y}_i-\boldsymbol{P}(\boldsymbol{X}_i,\boldsymbol{\beta})\right] + (\boldsymbol{\beta}-\boldsymbol{\mu})^T\boldsymbol{V}_{\mu}^{-1}(\boldsymbol{\beta}-\boldsymbol{\mu}) \quad (5)$$

where w_i denotes the weighting factor for the ith profile, $\boldsymbol{Y}_i$ denotes the ith SIMS profile, $\boldsymbol{P}(\boldsymbol{X}_i,\boldsymbol{\beta})$ denotes the simulated boron profile for the ith thermal annealing conditions $\boldsymbol{X}_i$ and the parameter estimates $\boldsymbol{\beta}$, and d is the total number of SIMS profiles. The TED model is coupled into the minimization

problem through $\boldsymbol{P}(\boldsymbol{X}_i, \boldsymbol{\beta})$, and simulated at each objective function evaluation. Here, the weighting factor is selected to equal the total number of data points in the corresponding SIMS profile. The covariance $\boldsymbol{V}_\epsilon$ is estimated by

$$\boldsymbol{V}_\epsilon = \left(\sum_{i=1}^{n} \frac{\left(Y_i - P(\boldsymbol{X},\boldsymbol{\beta})_i\right)^2}{n-1} \right) I \tag{6}$$

where $\boldsymbol{\beta}$ is fit to experimental data $\boldsymbol{Y}$.

In addition, several constraints must be included in the MAP estimation to arrive at physically meaningful estimates. One of the constraints is Ostwald ripening (Stolk et al., 1997), where clusters of smaller sizes dissociate to favor formation of larger clusters. This constrains clusters of larger sizes to have progressively higher dissolution activation energies. Furthermore, all activation energies are constrained to be positive due to the existence of energy barriers for these rate processes. Hence, the final form for the MAP estimation problem for the TED energetics is

$$\min_{\substack{\beta \\ \beta_i \geq 0 \\ E_2 \leq E_3 \\ E_3 \leq E_4 \\ E_4 \leq E_{large}}} \left\{ \sum_{i=1}^{d} \frac{1}{n_i} \left[\boldsymbol{Y}_i - \boldsymbol{P}(\boldsymbol{X}_i,\boldsymbol{\beta})\right]^T \boldsymbol{V}_{\epsilon_i}^{-1} \left[\boldsymbol{Y}_i - \boldsymbol{P}(\boldsymbol{X}_i,\boldsymbol{\beta})\right] + (\boldsymbol{\beta} - \boldsymbol{\mu})^T \boldsymbol{V}_\mu^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}) \right\} \tag{7}$$

where n_i is the number of data points in the ith SIMS profile.

Numerical Implementation

The constrained MAP estimation problem was numerically solved using the popular sequential simulation-optimization method, where the optimization algorithm calls the simulation algorithm, which is treated like any subroutine that produces an output (in this case, the MAP objective in Eq. 7 given inputs, that is, the parameters to be estimated). The CFSQP (Lawrence et al., 1997) optimization algorithm was used, which implements a modified sequential quadratic programming algorithm such that the constraints are satisfied at each iterate. CFSQP solves constrained minimization problems with possibly sequential objectives and sequential inequality and/or equality constraints, which include the MAP estimation problem. The optimization algorithm combines quadratic programming and Armijo-type line search to achieve superlinear convergence (Lawrence et al., 1997).

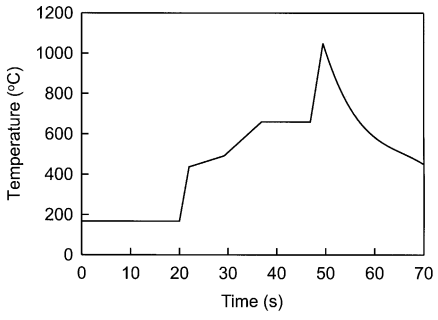

Figure 2. Typical rapid thermal anneal temperature program consisting of a stabilization step and a spike-anneal (that is, a fast linear heating step followed by a natural cooldown step).

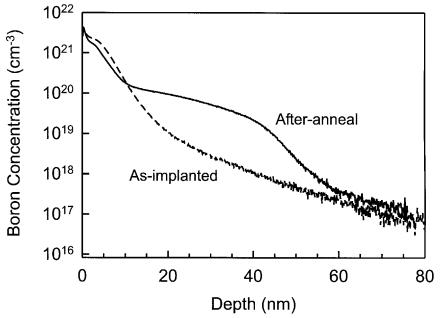

Figure 3. As-implanted and after-anneal boron SIMS profiles showing the profile spreading from transient enhanced diffusion corresponding to an RTA employing 150°C/s heating rate to an annealing temperature of 1,050°C.

The TED model was simulated using Alagator scripts running on the public domain software FLOOPS 2000 (Law and Cea, 1998). FLOOPS spatially discretizes the continuity equations to give an implicit system of ordinary differential equations (ODEs). The resulting ODEs are solved using the TR/BDF2 (Law and Cea, 1998) composite method for time integration, which combines the one-step trapezoidal rule (TR) and the multistep backward differentiation formula (BDF). The TR/BDF2 method has second-order accuracy and satisfies a numerical stability condition that is desirable for stiff differential equations (Bank et al., 1985). The FLOOPS simulations comprised essentially all of the computational cost in solving the MAP estimation for TED parameters.

The initial boron profile was the experimental as-implanted SIMS profile provided by International Sematech. The total boron was assumed to contain 20% substitutional boron and 80% interstitial boron, as suggested by experimental observations (Caturla et al., 1998; Kobayashi et al., 2001). The initial conditions for silicon interstitials presupposed with the "+1" model, wherein Si interstitial concentrations track the total boron concentration. Boundary conditions at the surface for all species assumed no flux (that is, $J_i|_{\text{surf}} = 0$) with no surface Fermi level pinning (Jung and Seebauer, 2000). A typical thermal annealing temperature trajectory is shown in Figure 2, with a ramp rate ranging from 100°C to 400°C/s to a maximum temperature around 1,000° to 1,200°C, followed by a natural cooling step. Figure 3 shows the as-implanted SIMS profile with a typical after-anneal profile showing the well-known transient enhanced diffusion of the boron profile with its increased junction depth.

Table 2. MAP Estimation for Different Parameter Sets

TED Parameters	A*	B**	C†
$E_{\text{diff, B}i^+}$	0.363	0.342	0.342
$E_{\text{diff, Si}_i^{+2}}$	††	††	0.720
E_{ko}	0.500	0.458	0.458
E_{ki}	0.510	0.508	0.508
E_{dis}	0.581	0.551	0.551
$E_{2,\text{B}}$	††	1.808	1.808
E_2	††	††	1.40
E_3	††	2.198	2.198
E_4	††	2.995	2.995
E_{large}	††	††	3.70
$E_{\text{large,mix}}$	††	††	3.50

*Most sensitive parameters.
**Most and moderately sensitive parameters.
†All parameters.
††Parameter value is fixed to be equal to the ML estimate.

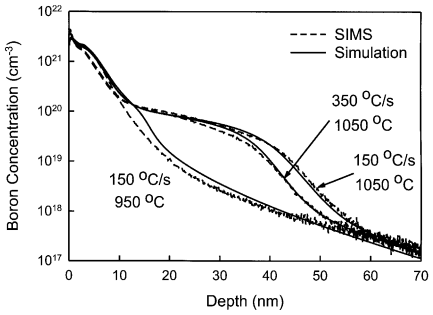

Figure 5. After-anneal SIMS and simulation profiles using MAP estimates of TED energetics (see Table 3) for various heating rates and annealing temperatures.

Results and Discussion

A preliminary study using one SIMS profile elucidated the dependence of the MAP estimates on different combinations of TED parameters. Here, the MAP estimation utilized the SIMS profile employing RTA with a heating rate of 150°C/s and a maximum temperature of 1,050°C. Table 2 presents the MAP estimates for three sets of activation energies that were estimated: (A) the most sensitive, (B) the most to moderately sensitive, and (C) all activation energies in the TED model. Figure 4 shows the simulation profiles associated with the MAP and the prior estimates determined using ML. Negligible improvement in the profile fitting resulted from using the MAP estimates in set A as opposed to the ML parameters. This was expected from the closeness of the set A MAP parameters to the ML estimates, which suggests a lack of informativeness of the SIMS profile on these parameters. However, the simulation profiles for the set B and set C MAP parameters provide an improved fit over the set A and ML parameters. This improvement indicates the importance of boron and silicon clustering as interstitial sinks and sources during the annealing process (the difference between sets B and C compared to set A are the parameters for boron and silicon clustering, as shown in Table 2). The inclusion of the additional insensitive parameters in set C in MAP estimation provided no improvement over set B, in agreement with parameter sensitivity results (Gunawan et al., 2002a). None of the parameter sets match the profile close to the surface ($<$15 nm). The lack of fit in this region did not arise from inaccurate parameter estimates, but from the inadequacy of the no-flux surface boundary condition in capturing dose loss during annealing. This suggests the importance of undergoing research on surface effects, such as dose-loss kinetics and Fermi-level pinning, to improve the TED model.

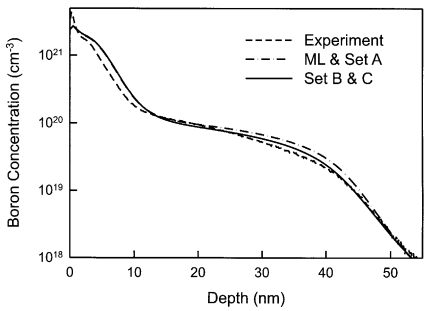

Figure 4. Experimental and simulation profiles using ML and MAP parameter estimates in Tables 1 and 2.

The SIMS profile corresponds to a spike-annealing at temperature 1,050°C with a ramp-up rate of 150°C/s.

The preliminary study just discussed suggested that the full MAP estimation should focus on the parameters in set B, because of the lack of informativeness of the SIMS profiles for parameter combinations in sets A and C. The complete MAP estimation used three sets of SIMS profiles corresponding to RTAs with different ramp rates and maximum temperatures (150°C/s to 1,050°C, 350°C/s to 1,050°C, and 150°C/s to 950°C) provided by International Sematech. Table 3 presents the MAP estimates along with the confidence intervals and standard deviations. Figure 5 compares the simulation profiles using the MAP estimates with the SIMS profiles, which demonstrates the ability of the TED model to match experimental observations. Again, there was lack of profile fit close to the surface due to the aforementioned reasons. These differences are more pronounced for the profile with a maximum temperature of 950°C, because little TED occurred, and, thus, surface effects had a greater effect. Furthermore, comparison of the simulation profiles using the MAP and ML estimates (see Figures 1 and 5) showed the improvement in the profile fitting of MAP parameter estimates compared to ML *a priori* parameter estimates, especially for the profile with a ramp rate of 350°C/s.

Figure 6 shows good agreement between simulations of the TED model using MAP estimates and published experimental sheet resistance–junction depth pairs (Agarwal et al., 1999) employing various spike annealing programs, implant energies, and dopant doses. Figure 6 also shows the "Sematech curve" (Murto, 1999), which is the phenomenological locus of points commonly used to map the "best" pairs of junction depth and sheet resistance typically obtained by conventional spike annealing. The simulations employed ranges of heating rates between 50 and 400°C/s, cooling rates between 50 and 150°C, and maximum temperatures between 950 and 1,200°C. To our knowledge, Figure 6 represents the first published agreement between model simulations of TED and non-SIMS experimental results. The agreement indicates that the TED model and MAP estimates provide a good predictive capabil-

Table 3. Maximum *A Posteriori* Estimates of Transient Enhanced Diffusion Energetics

TED Parameters	MAP Est.	95% Conf. Int. (Std. Dev. Est.*)
$E_{\text{diff, B}_i^+}$	0.359	0.008 (0.004)
E_{ko}	0.408	0.015 (0.008)
E_{ki}	0.458	0.015 (0.008)
E_{dis}	0.575	0.004 (0.002)
$E_{2,\text{B}}$	1.788	0.024 (0.012)
E_3	2.192	0.023 (0.012)
E_4	3.055	0.004 (0.002)

*The standard deviation is computed from the confidence interval using z statistics (Mann, 1995).

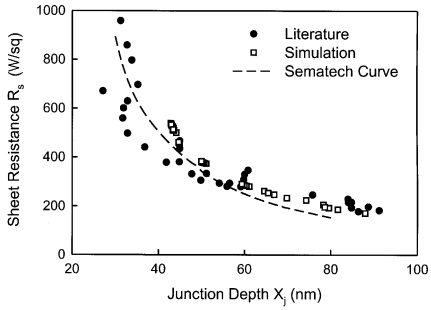

Figure 6. Junction depth–sheet resistance pairs: published experimental works vs. TED simulations employing various heating and cooling rates, and annealing temperatures.

Sematech curve summarized the sheet resistance and junction depth data in experimental studies performed by International Sematech.

ity for designing optimal annealing programs to achieve the desired junction depth and sheet resistance.

The estimated covariance of the *a posteriori* parameters (see Eq. A18 in the Appendix) was

$$V_{\beta^*} = \begin{bmatrix} E_{diff,\text{B}_i^+} & E_{ko} & E_{ki} & E_{dis} & E_3 & E_4 & E_{2,B} \\ 4.42\times10^{-5} & -1.33\times10^{-4} & -9.64\times10^{-5} & 1.66\times10^{-5} & -3.00\times10^{-5} & 3.34\times10^{-6} & 5.13\times10^{-5} \\ -1.33\times10^{-4} & 4.50\times10^{-4} & 3.35\times10^{-4} & -4.73\times10^{-5} & 1.34\times10^{-4} & -4.19\times10^{-6} & -2.12\times10^{-4} \\ -9.64\times10^{-5} & 3.35\times10^{-4} & 2.52\times10^{-4} & -3.44\times10^{-5} & 1.03\times10^{-4} & -1.67\times10^{-6} & -1.62\times10^{-4} \\ 1.66\times10^{-5} & -4.73\times10^{-5} & -3.44\times10^{-5} & 7.35\times10^{-6} & -5.10\times10^{-6} & 1.44\times10^{-6} & 1.28\times10^{-5} \\ 3.00\times10^{-5} & 1.34\times10^{-4} & 1.03\times10^{-4} & -5.10\times10^{-6} & 8.94\times10^{-5} & 4.86\times10^{-7} & -1.11\times10^{-4} \\ 3.34\times10^{-6} & -4.19\times10^{-6} & -1.67\times10^{-6} & 1.44\times10^{-6} & 4.86\times10^{-7} & 1.43\times10^{-6} & -7.44\times10^{-7} \\ 5.13\times10^{-5} & -2.12\times10^{-4} & -1.62\times10^{-4} & 1.28\times10^{-5} & -1.11\times10^{-4} & -7.44\times10^{-7} & 1.49\times10^{-4} \end{bmatrix} \quad (8)$$

The estimated covariance shows strong collinearity among many of the parameters and indicates that the confidence interval estimates in Table 3 may not be accurate representations of the confidence region for the parameter estimates. A linear correlation matrix r measures the degree of linear dependence between a pair of parameters where a magnitude of 1 describes perfect correlation and a value of 0 suggests no correlation. The correlation matrix consists of correlation coefficients whose values are given by (Beck and Arnold, 1977)

$$r_{ij} = V_{ij}\left(V_{ii}V_{jj}\right)^{-1/2} \quad (9)$$

Computation of the correlation matrix for V_{β^*} gave

which more clearly indicates the strong collinearity among the parameters. The most apparent collinearities are those associated with boron diffusion and activation parameters. Another strong collinearity arises between pure boron and size 3 cluster dissolution energetics. The strong collinearity is caused by the lack of information in the SIMS profiles to simultaneously identify these parameters due to poor experimental design and lack of dynamical data (SIMS profiles collected at different times during the annealing process). Moreover, parameter sensitivity analysis (Gunawan et al., 2002a) showed that the perturbations in these parameters exert similar changes in the profile shape, which makes independent identification of each parameter challenging. This result motivates the application of the optimal model-based experimental design (Beck and Arnold, 1977; Gunawan et al., 2002b) to maximize the informativeness of the data, and to carry out model selection among competing hypothetical mechanisms.

$$r = \begin{bmatrix} 1 & -0.95 & -0.91 & 0.92 & -0.48 & 0.42 & 0.63 \\ -0.95 & 1 & 0.99 & -0.82 & 0.67 & -0.17 & -0.82 \\ -0.91 & 0.99 & 1 & -0.80 & 0.69 & -0.09 & -0.84 \\ 0.92 & -0.82 & -0.80 & 1 & -0.20 & 0.44 & 0.39 \\ -0.48 & 0.67 & 0.69 & -0.20 & 1 & 0.04 & -0.96 \\ 0.42 & -0.17 & -0.09 & 0.44 & 0.04 & 1 & -0.05 \\ 0.63 & -0.82 & -0.84 & 0.39 & -0.96 & -0.05 & 1 \end{bmatrix} \quad (10)$$

Despite the strong collinearity, it is useful to compare the MAP and ML parameter estimates and their standard deviations to elucidate the informativeness of the SIMS data with respect to each parameter, and the quality of the ML estimates. The standard deviations of MAP parameters were smaller than those of ML estimates (as expected from Eq. A18 in the Appendix). In particular, there was a significant reduction in the standard deviation associated with the activation energy of $(B_s - Si_i)$ dissociation into substitutional boron E_{dis} (in correlation with k_{dis} of Eq. A8 in the Appendix). This indicates that the SIMS profiles contain more information about this parameter than the others. Although the standard deviation for the size-3 cluster dissolution energy E_3 is of similar magnitude to E_{dis}, lack of prior information on its uncertainty prohibited such a conclusion. Most of the MAP and ML intervals overlap each other with the exception of the dissociation energy of pure boron cluster $E_{2,B}$, which indicates that the *a priori* information for this parameter may be inaccurate. This motivates further experimental and/or density functional theory calculations for estimating this parameter, to provide an independent confirmation of the MAP parameter estimate determined here.

Conclusions

Although the number of parameters associated with TED energetics is large and the information content from after-anneal SIMS boron profiles is limited, MAP estimation is able to combine *a priori* information with the SIMS profiles to give a nonsingular and well-posed estimation of the parameters and their accuracies. The results also suggest that MAP estimation provides estimates of the TED parameters with greater accuracy than ML estimates obtained from individual focused experimental studies and density functional theory calculations. The strong collinearity in the *a posteriori* parameter covariance matrix suggests poor experimental design and the need to employ optimal model-based experimental design (Beck and Arnold, 1977). Comparison of MAP and ML simulations confirms the importance of surface effects, especially dose loss, during TED.

Acknowledgment

Support from International Sematech is acknowledged.

Literature Cited

Agarwal, A., H.-J. Gossmann, and A. T. Fiory, "Effect of Ramp Rates During Rapid Thermal Annealing of Ion Implanted Boron for Formation of Ultra-Shallow Junctions," *J. Electronic Materials*, **28**, 1333 (1999).

Agarwal, A., L. Pelaz, and H. H. Vuong, eds., *Silicon Front-End Processing — Physics & Technology of Dopant-Defect Interactions II*, Vol. 610, Materials Research Society, Warrendale, PA (2000).

Bahl, L. R., R. Bakis, F. Jelinek, and R. L. Mercer, "Language-Model/Acoustic Channel Balance Mechanism," *IBM Technical Disclosure Bulletin*, 3464 (1980).

Bank, R. E., W. M. Coughran, Jr., W. Fichtner, E. H. Grosse, D. J. Rose, and R. K. Smith, "Transient Simulation of Silicon Devices and Circuits," *IEEE Trans. Elecron Devices*, **ED-32**, 1992 (1985).

Beck, J. V., and K. J. Arnold, *Parameter Estimation in Engineering and Science*, Wiley, New York (1977).

Bennett, D. J., and T. E. Price, "Transient Enhanced Diffusion During Post-Implant Annealing of Silicon," *Semiconductor Science & Technology*, **9**, 1535 (1994).

Caturla, M. J., M. Foad, and T. Daiz de la Rubia, "The Effect of Ramp Rate and Annealing Temperature on Boron Transient Diffusion in Implanted Silicon: Kinetic Monte Carlo Simulations," *Proc. Int. Conf. on Ion Implantation Technology*, Kyoto, Japan, p. 1018 (1998).

Chang, I. S., and R.-H. Park, "Segmentation Based on Fusion of Range and Intensity Images Using Robust Trimmed Methods," *Pattern Recognition*, **34**, 1951 (2001).

Collart, E. J. H., 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, "Cluster Formation During Annealing of Ultra-Low-Energy Boron-Implanted Silicon," *J. Vacuum Science & Technology*, **18**, 435 (2000).

Cowern, N. E. B., K. T. F. Janssen, G. F. A. van de Walle, and D. J. Gravesteijn, "Impurity Diffusion via an Intermediate Species: The B-Si System," *Phys. Rev. Lett.*, **65**, 2434 (1990).

Eaglesham, D., "Dopants, Defects and Diffusion," *Phys. World*, **8**, 41 (1995).

Gunawan, R., M. Y. L. Jung, R. D. Braatz, and E. G. Seebauer, "Systems Analysis Applied to Modeling Dopant Activation and TED in Rapid Thermal Annealing," *Proc. IEEE Int. Conf. of Advanced Thermal Processing of Semiconductors*, Vancouver, Canada, p. 107 (2002a).

Gunawan, R., D. L. Ma, M. Fujiwara, and R. D. Braatz, "Identification of Kinetic Parameters in Multidimensional Crystallization Processes," *Int. J. Mod. Phys. B*, **16**, 367 (2002b).

Haynes, T. E., D. J. Eaglesham, P. A. Stolk, H. J. Gossmann, D. C. Jacobson, and J. M. Poate, "Interactions of Ion-Implantation-Induced Interstitials with Boron at High Concentrations in Silicon," *Appl. Phys. Lett.*, **69**, 1376 (1996).

Jung, M. Y. L., and E. G. Seebauer, "New Physics for Modeling Transient Enhanced Diffusion in RTP," *Proc. Rapid Thermal and Other Short-Time Processing Tech. II*, ECS, Toronto, Canada, p. 15 (2000).

Kobayashi, H., I. Nomachi, S. Kusanagi, and F. Nishiyama, "Lattice Site Location of Ultra-Shallow Implanted B in Si Using Ion Beam Analysis," *Si Front-End Processing — Physics & Technology of Dopant-Defect Interactions III*, K. S. Jones, M. D. Giles, P. Stolk, J. Matsuo, and E. C. Jones, eds., Vol. 669, Materials Research Society, Inc., Warrendale, PA, p. J5.3 (2001).

Laidler, K. J., *Chemical Kinetics*, 3rd ed., Harper & Row, New York (1987).

Lamberg, L., K. Muinonen, J. Ylonen, and K. Lumme, "Spectral Estimation of Gaussian Random Circles and Spheres," *J. Comput. Appl. Math.*, **136**, 109 (2001).

Law, M. E., and S. M. Cea, "Continuum Based Modeling of Silicon Integrated Circuit Processing: An Object Oriented Approach," *Comput. Mat. Sci.*, **12**, 289 (1998).

Law, M. E., H. Park, and P. Novell, "Theory of Dopant Diffusion Assuming Nondilute Concentrations of Dopant-Defect Pairs," *Appl. Phys. Lett.*, **59**, 3488 (1991).

Lawrence, C., J. L. Zhou, and A. L. Tits, "User's Guide for CFSQP Version 2.5: A C Code for Solving (Large Scale) Constrained Nonlinear (Minimax) Optimization Problems, Generating Iterates Satisfying All Inequality Constraints," Software Manual, Univ. of Maryland, College Park, MD (1997).

Lerch, W., M. Glück, N. A. Stolwijk, H. Walk, M. Schäfer, S. D. Marcus, D. F. Downey, and J. W. Chow, "Boron Ultrashallow Junction Formation in Silicon by Low-Energy Implantation and Rapid Thermal Annealing in Inert and Oxidizing Ambient," *J. Electrochem. Soc.*, **146**, 2670 (1999).

Mann, P. S., *Introductory Statistics*, Wiley, New York (1995).

Murto, B., "Recent Advances and Continuing Challenges in Ultra-Shallow Junctions," *Proc. National Implant Users Meeting*, Austin, TX (1999).

Seebauer, E. G., and M. Y. L. Jung, "Surface Diffusion of Adsorbates on Metals, Alloys, Oxides, and Semiconductors," *Landolt-Börnstein Numerical Data and Functional Relationships: Adsorbed Layers on Surfaces*, H. P. Bonzel, ed., Vol. III/42A, Springer-Verlag, New York (2001).

Stolk, P. A., H. J. Gossmann, D. J. Eaglesham, D. C. Jacobson, C. S. Rafferty, G. H. Gilmer, M. Jaraiz, J. M. Poate, H. S. Luftman, and T. E. Haynes, "Physical Mechanisms of Transient-Enhanced Dopant Diffusion in Ion-Implanted Silicon," *J. Appl. Phys.*, **81**, 6031 (1997).

Van Vechten, J. A., "Activation Enthalpy of Recombination-Enhanced Vacancy Migration in Si," *Phys. Rev. B (Condensed Matter)*, **38**, 9913 (1988).

Van Vechten, J. A., and C. D. Thurmond, "Entropy of Ionisation and Temperature Variation of Ionisation Levels of Defects in Semiconductors," *Phys. Rev. B (Solid State)*, **14**, 3539 (1976).
Varma, A., M. Morbidelli, and H. Wu, *Parametric Sensitivity in Chemical Systems*, Cambridge Univ. Press, Melbourne, Australia (1999).
Wang, Z., and E. G. Seebauer, "Estimating Pre-Exponential Factors for Desorption from Semiconductors: Consequences for *A Priori* Process Modeling," *Appl. Surf. Sci.*, **181**, 111 (2001).
Xiao, G., M. Brady, J. A. Noble, and Y. Zhang, "Segmentation of Ultrasound B-Mode Images with Intensity Inhomogeneity Correction," *IEEE Trans. Med. Imaging*, **21**, 48 (2002).
Yariv, A., *An Introduction to Theory and Applications of Quantum Mechanics*, Wiley, New York (1982).
Zarnich, R. E., K. L. Bell, and H. L. Van Trees, "A Unified Method for Measurement and Tracking of Contacts from an Array of Sensors," *IEEE Trans. Signal Process.*, **49**, 2950 (2001).
Zhang, L. H., K. S. Jones, P. H. Chi, and D. S. Simons, "Transient Enhanced Diffusion without (311) Defects in Low Energy B^+-Implanted Silicon," *Appl. Phys. Lett.,* **67**, 2025 (1995).

Appendix

Supplemental TED Model

The flux J_i includes a Fickian diffusion term and a drift motion term from the electric field effect on the charged species

$$J_i = -D_i \frac{\partial^2 N_i}{\partial x^2} + \gamma_i \mu_i N_i E(x) \tag{A1}$$

where D_i denotes the diffusivity and $E(x)$ is the electric field. The mobility μ_i follows the Einstein relation

$$\mu_i = \frac{qD_i}{kT} \tag{A2}$$

where q is the electron charge, k is the Boltzmann constant, and T is the temperature. The term γ_i describes the net charge according to

$$\gamma_i = \sum_j z_j \gamma_{i,z_j} \tag{A3}$$

where the z_j are the charge states (that is, $+2$, 0, -1, and so on) and γ_{i,z_j} is the fraction of species i with charge z_j according to the Fermi Dirac statistics (Yariv, 1982). For example, a species C that possesses a charge distribution of 30% C^{+2} and 70% C^- has a net charge of -0.1. Rewriting the electric field as a function of the potential ψ

$$E(x) = -\frac{\partial \psi}{\partial x} \tag{A4}$$

and assuming that the diffusivity D_i is invariant of the specie charges, the continuity equation for mobile species during TED can be written as

$$\frac{\partial N_i}{\partial t} = D_i \left[\frac{\partial^2 N_i}{\partial x^2} + \frac{q}{kT} \left(N_i \gamma_i \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial \psi}{\partial x} \frac{\partial}{\partial x} (N_i \gamma_i) \right) \right] + G_i \tag{A5}$$

For immobile species, the first term in the righthand side of the continuity equation drops out ($D_i = 0$), leaving only G_i.

The clusters that form during ion implantation and thermal annealing (Collart et al., 2000; Stolk et al., 1997) act as interstitial traps (at a lower temperature) as well as sources (at a higher temperature). Evidence exists to support the formation of pure boron (Collart et al., 2000), pure silicon (Stolk et al., 1997), and mixed boron–silicon clusters (Haynes et al., 1996) during TED. The formation and dissolution of the pure interstitial clusters follow the reactions

$$I + I \underset{d}{\overset{f}{\rightleftharpoons}} I_2$$

$$I_{m-1} + I \underset{d}{\overset{g}{\rightleftharpoons}} I_m$$

where I denotes the interstitial (boron and silicon) and the indices m denote the size of the cluster. The reaction rates for the formation, growth, and dissociation of pure clusters follow

$$r_f = 8\pi a D_I [I]^2$$

$$r_g = 4\pi a D_I [I][I_{m-1}]$$

$$r_d = k_m [I_m] \tag{A6}$$

where a is the capture radius ($a = 2.7$ Å) and k_m is the dissociation rate constant for size m clusters. The cluster formation rate is assumed to be diffusion limited by the reactant, in agreement with much of the literature (see, for example, Laidler, 1987). The dissolution rate assumes the existence of an energy barrier as the rate-limiting step such that the rate constant follows the Arrhenius law.

The mixed boron–silicon cluster formation and dissolution are given by

$$\mathrm{B}_{p-1}\mathrm{Si}_q + \mathrm{B}_i \underset{\hat{d}}{\overset{\hat{g}}{\rightleftharpoons}} \mathrm{B}_p\mathrm{Si}_q$$

$$\mathrm{B}_p\mathrm{Si}_{q-1} + \mathrm{Si}_i \underset{\hat{d}}{\overset{\hat{g}}{\rightleftharpoons}} \mathrm{B}_p\mathrm{Si}_q$$

where p, q are integers greater than or equal to 1. The reaction rates of mixed clustering follows similar rate equations as in pure interstitial clusters (see Eq. A6), in accord to similar arguments above

$$r_{\hat{g}} = 4\pi a \left(D_{\mathrm{B}_i} \right) [\mathrm{B}_i] \left[\mathrm{B}_{p-1}\mathrm{Si}_q \right] \quad \text{or} \quad 4\pi a \left(D_{\mathrm{Si}_i} \right) [\mathrm{Si}_i] \left[\mathrm{B}_p\mathrm{Si}_{q-1} \right]$$

$$r_{\hat{d}} = k_{p+q} \left[\mathrm{B}_p\mathrm{Si}_q \right] \tag{A7}$$

Due to the limitation in the number of the differential equations that can be solved by FLOOPS, the sizes of interstitial clusters are limited to two for pure boron and five for pure silicon and mixed boron–silicon clusters. This should not impose a severe limitation on the TED model, as there is no evidence for the formation of large clusters in sub-keV implants (Zhang et al., 1995).

The boron activation reaction plays a very important role in TED as a medium of transformation between immobile substitutional boron and mobile interstitial boron. The intermediate species B_s–Si_i also act as nucleation centers for the mixed boron–silicon cluster. The activation occurs according to

$$B_s + Si_i \underset{\text{dis}}{\overset{\text{assoc}}{\rightleftharpoons}} (B_s - Si_i) \overset{\text{Kick-out}}{\underset{\text{Kick-in}}{\overset{ko}{\underset{ki}{\rightleftharpoons}}}} B_i + Si_s$$

The reaction rates are given by

$$\begin{aligned} r_{\text{assoc}} &= k_{\text{assoc}}[B_s][Si_i] \\ r_{\text{dis}} &= k_{\text{dis}}[B_s - Si_i] \\ r_{ki} &= k_{ki}[B_i][Si_i] \\ r_{ko} &= k_{ko}[B_s - Si_i] \end{aligned} \tag{A8}$$

where the reaction rate constants (that is, k_{ko}, k_{ki}, k_{assoc}, and k_{dis}) are again assumed to follow the Arrhenius law.

Maximum A Posteriori estimation

Let $f(\cdot)$ denote the probability density function and $f(A|B)$ denote the conditional probability function of the event A given the event B. A vector of random variables $\boldsymbol{x} \in \mathbb{R}^m$ follows a normal distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^m$ and variance $\boldsymbol{V} \in \mathbb{R}^{m \times m}$, denoted by $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{V})$, if

$$f(\boldsymbol{x}) = (2\pi)^{-m/2} |\boldsymbol{V}|^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{V}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right] \tag{A9}$$

where $|\boldsymbol{A}|$ denotes the determinant of matrix $\boldsymbol{A}$. Bayes' theorem describes the *a posteriori* distribution as a function of the *a priori* distribution and experimental data (Beck and Arnold, 1977)

$$f(\boldsymbol{\beta}|\boldsymbol{Y}) = \frac{f(\boldsymbol{Y}|\boldsymbol{\beta}) f(\boldsymbol{\beta})}{f(\boldsymbol{Y})} \tag{A10}$$

The variables $\boldsymbol{\beta}$ and $\boldsymbol{Y}$ denote the vectors of parameter estimates and experimental observations, respectively, and $f(\boldsymbol{\beta}|\boldsymbol{Y})$ denotes the *a posteriori* distribution for the parameter estimates given the data $\boldsymbol{Y}$. The MAP parameters $\boldsymbol{\beta}^*$ are the modes of $f(\boldsymbol{\beta}|\boldsymbol{Y})$

$$\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\beta}|\boldsymbol{Y})\big|_{\boldsymbol{\beta}^*} = 0 \quad \text{or} \quad \left.\frac{\partial f(\boldsymbol{\beta}|\boldsymbol{Y})}{\partial \beta_j}\right|_{\boldsymbol{\beta}^*} = 0 \quad \text{for} \quad j = 1, 2, 3, \ldots, p \tag{A11}$$

where p is the number of parameters.

Experimental errors are assumed to be additive with a normal distribution of zero mean and variance $\boldsymbol{V}_\epsilon$

$$\boldsymbol{Y} = \boldsymbol{P}(\boldsymbol{X}, \boldsymbol{\beta}) + \boldsymbol{\epsilon} \tag{A12}$$

where $\boldsymbol{P}(\boldsymbol{X}, \boldsymbol{\beta})$ describes the process model, $\boldsymbol{X}$ denote the independent variables, and the error $\boldsymbol{\epsilon} \sim N(0, \boldsymbol{V}_\epsilon)$. Under this condition and assuming no error in $\boldsymbol{X}$, the function $f(\boldsymbol{Y}|\boldsymbol{\beta})$ describes the probability of obtaining the experimental data $\boldsymbol{Y}$ given the parameters $\boldsymbol{\beta}$

$$f(\boldsymbol{Y}|\boldsymbol{\beta}) = (2\pi)^{-n/2} |\boldsymbol{V}_\epsilon|^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{Y} - \boldsymbol{P})^T \boldsymbol{V}_\epsilon^{-1} (\boldsymbol{Y} - \boldsymbol{P}) \right] \tag{A13}$$

where n is the total number of measurements. The *a priori* distribution for the parameters $\boldsymbol{\beta}$ is assumed to follow a normal distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{V}_\mu$

$$f(\boldsymbol{\beta}) = (2\pi)^{-p/2} |\boldsymbol{V}_\mu|^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{\beta} - \boldsymbol{\mu})^T \boldsymbol{V}_\mu^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}) \right] \tag{A14}$$

The parameters and measurements are assumed to be uncorrelated that is $\text{cov}(\boldsymbol{\beta}, \boldsymbol{\epsilon}) = 0$.

Using the fact that the maximizer of a function also maximizes its logarithm, reformulation of the MAP estimation problem in Eq. A11 gives

$$\left[\nabla_{\boldsymbol{\beta}} \ln f(\boldsymbol{Y}|\boldsymbol{\beta}) + \nabla_{\boldsymbol{\beta}} \ln f(\boldsymbol{\beta}) \right]\big|_{\boldsymbol{\beta}^*} = 0 \tag{A15}$$

Under the aforementioned assumptions, substitution of Eqs. A13 and A14 into Eq. A15 transforms the MAP estimation into the minimization problem

$$\min_{\boldsymbol{\beta}} S_{\text{MAP}} = \min_{\boldsymbol{\beta}} \left\{ \ln|\boldsymbol{V}_\epsilon| + \ln|\boldsymbol{V}_\mu| + (\boldsymbol{Y} - \boldsymbol{P})^T \boldsymbol{V}_\epsilon^{-1} (\boldsymbol{Y} - \boldsymbol{P}) + (\boldsymbol{\beta} - \boldsymbol{\mu})^T \boldsymbol{V}_\mu^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}) \right\} \tag{A16}$$

If $\boldsymbol{V}_\epsilon$ and $\boldsymbol{V}_\mu$ are independent of $\boldsymbol{\beta}$, the MAP estimation problem further reduces to

$$\min_{\boldsymbol{\beta}} \left\{ (\boldsymbol{Y} - \boldsymbol{P})^T \boldsymbol{V}_\epsilon^{-1} (\boldsymbol{Y} - \boldsymbol{P}) + (\boldsymbol{\beta} - \boldsymbol{\mu})^T \boldsymbol{V}_\mu^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}) \right\} \tag{A17}$$

The first term of the objective function just given is also known as the likelihood function, and the MAP estimation reduces to ML estimation in the absence of *a priori* information (that is, $\boldsymbol{V}_\mu^{-1} = 0$). Using linearization of the model, an estimate of the covariance of $(\boldsymbol{\beta}^* - \boldsymbol{\beta}_{\text{true}})$ is given by (Beck and Arnold, 1977)

$$\text{cov}(\boldsymbol{\beta}^* - \boldsymbol{\beta}_{\text{true}}) = \boldsymbol{V}_{\beta^*} \approx \left(\boldsymbol{S}^T \boldsymbol{V}_\epsilon^{-1} \boldsymbol{S} + \boldsymbol{V}_\mu^{-1} \right)^{-1} \tag{A18}$$

where $\boldsymbol{\beta}_{\text{true}}$ denotes the true parameters and $\boldsymbol{S}$ denotes the sensitivity of the model $\boldsymbol{P}$ with respect to the parameters $\boldsymbol{\beta}$. Since computation of the sensitivities for the TED model or complex systems is expensive, the finite difference method (Varma et al., 1999) was used to estimate the sensitivities. If additional experimental data become available, then the MAP estimator $\boldsymbol{\beta}^*$ and the covariance estimate cov$(\boldsymbol{\beta}^* - \boldsymbol{\beta}_{\text{true}})$ would provide the *a priori* distribution for further MAP estimation to improve parameter estimates.

A hyperellipsoidal confidence region that quantifies the accuracy of the parameters is given by

$$E_{\boldsymbol{\beta}} = \left\{ \boldsymbol{\beta} : (\boldsymbol{\beta} - \boldsymbol{\beta}_{\text{true}})^T \boldsymbol{V}_{\boldsymbol{\beta}^*}^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}_{\text{true}}) \le \chi_\alpha^2(p) \right\} \quad \text{(A19)}$$

where α denotes the confidence level and $\chi_\alpha^2(p)$ denotes the chi-squared distribution with p degrees of freedom. The confidence intervals associated with the aforementioned confidence region are

$$-\left| \left(\chi_\alpha^2(p) V_{\beta^*,ii} \right)^{1/2} \right| \le (\beta_i^* - \beta_{\text{true,i}}) \le \left| \left(\chi_\alpha^2(p) V_{\beta^*,ii} \right)^{1/2} \right| \quad \text{(A20)}$$

where $V_{\beta^*,ii}$ denotes the ith diagonal element the covariance matrix $\boldsymbol{V}_{\beta^*}$. The confidence intervals are an accurate representation for the hyperellipsoidal confidence region only when the off-diagonal elements are much smaller than the diagonal elements.

Manuscript received Oct. 25, 2002, and revision received Feb. 19, 2003.