Obtaining Model-Independent Growth Rates from Experimental Data of Dry Thermal Oxidation of Silicon

Y. Leong Yeow
Dept. of Chemical and Biomolecular Engineering, The University of Melbourne, Victoria 3010, Australia

Jong-Leng Liow
School of Engineering and Information Technology, UNSW@ADFA, Canberra ACT 2600, Australia

Yee-Kwong Leong
School of Mechanical and Chemical Engineering, The University of Western Australia, Crawley, WA 6009, Australia

DOI 10.1002/aic.14375
Published online January 29, 2014 in Wiley Online Library (wileyonlinelibrary.com)

Empirical time-oxide layer thickness data of dry thermal oxidation of silicon were converted numerically into instantaneous growth rates by Tikhonov regularization. These growth rates are independent of any assumed kinetics of oxidation or functional form of the original data and can, therefore, be used in model testing of this important industrial process. A number of numerical issues in the Tikhonov regularization computation, for example, the expected monotonic decrease in growth rate with time and the large time span of some of the datasets, are addressed. Numerical results indicate that kinetic data presented in the form of growth rate-oxide layer thickness plots are more sensitive than the traditional time-oxide thickness plots for use in model testing and for parameter determination. The advantages of this new approach are demonstrated by three case studies covering the complete oxide thickness spectrum of interest to the microelectronic industry. © 2014 American Institute of Chemical Engineers *AIChE J*, 60: 1810–1820, 2014
Keywords: silicon oxidation, Deal–Grove model, integral equation, numerical differentiation, Tikhonov regularization

Introduction

Controlled growth of a layer of silicon dioxide (SiO_2) on silicon substrates is a key step in the fabrication of microelectronic devices.[1,2] A widely adopted method of building up such a layer is by dry thermal oxidation. In this process, silicon substrates are exposed to an oxygen atmosphere from as low as 300°C to as high as 1200°C. The stoichiometry of the oxidation reaction is represented by

$$Si + O_2 \rightarrow SiO_2 \quad (1)$$

This reaction is generally performed in tubular quartz reactors placed in a resistance-heated furnace. Silicon substrates are placed in slotted quartz carriers and pushed into the reactor. Depending on the final application, the thickness of the SiO_2 layer required varies from about a few nm to over 50 nm. To ensure the required thickness is formed on the substrate, the conditions of the reaction, such as furnace temperature, oxygen pressure, and oxidation time, have to be carefully controlled. Exposed silicon reacts with oxygen to form a thin oxide film. For the reaction to continue, the oxygen has to diffuse through the oxide layer to reach the unreacted silicon. Thus, the instantaneous rate of growth of the oxide layer $dh(t)/dt$ depends on the diffusivity of oxygen through the oxide layer, the thickness of the layer $h(t)$, and the kinetics of the oxidation reaction (the explicit dependence on t in $h(t)$ will be suppressed where this does not lead to ambiguity).

In view of its importance to the microelectronics industry, the growth of the oxide layer has been investigated over a range of temperatures, oxygen partial pressures, and for different silicon crystal orientations. Numerous sets of tabulated time vs. thickness, t-h, data can be found in the open literature. Numerous models have also been proposed to explain the observed kinetics of oxide growth. Most of these can be regarded as modifications of the original phenomenological model of Deal and Grove.[3] These authors assumed that the flux of O_2 diffusing through the oxide layer is described by the pseudo steady-state diffusion equation and that the kinetics of the oxidation reaction at the SiO_2-Si interface is first order in the interfacial O_2 concentration. Equating the diffusion flux of O_2 through the oxide layer to the rate of consumption of oxygen at the interface they obtained the following relationship between h and t[1,2]

$$h(t)^2 + Ah(t) = B(t+\tau) \quad (2)$$

where A, B, and τ are the parameters of the model. Their numerical values depend on the diffusivity of O_2, the surface reaction rate constant, the concentration of O_2 at the exposed SiO_2 surface, and the thickness of any native oxide layer that may be present on the silicon substrate at the start of the thermal oxidation process. Equation 2 is now known as

Correspondence concerning this article should be addressed to Y. L. Yeow at yly@unimelb.edu.au.

the Deal–Grove linear-parabolic model of the oxidation process where B is referred to as the parabolic rate constant and B/A, the linear rate constant.[1,2]

In dry thermal oxidation, the linear-parabolic model is able to describe the increase in $h(t)$ with t accurately for $h(t) > 30$ nm. For $h(t) < 30$ nm, this model consistently underestimates $h(t)$.[1] A large number of modifications to the assumptions made by Deal and Grove have been put forward to improve its predictive performance. For example, it has been proposed that oxygen is transported as atomic or ionic oxygen O instead of molecular O_2 through the oxide layer.[4] However, there are claims by subsequent authors that this is not supported by experimental evidence and/or theoretical analyses.[5] A number of investigators suggested that the oxidation reaction may not be localized at the SiO_2-Si interface but is spread out over a region of finite thickness or even over the entire oxide layer.[6–8] Other models replace the simple first-order surface reaction kinetics with kinetics of increasing complexity. For example, it has been proposed that electrostatic charges close to the SiO_2-Si interface may play a role in controlling the rate of the oxidation reaction.[9] Because of the difference in the specific volume of Si and SiO_2, stress is generated at the interface and various investigations suggest that the stress or strain may alter the kinetics of the oxidation reaction or the diffusivity of O_2 and may be responsible for the accelerated initial growth.[10–14] The fundamental mechanism or explanation, at the microscopic/atomic level, of many of these modifications remains an open question. The oxidation reaction is known to be exothermic, hence thermal effects may, therefore, be important and may be responsible for the observed accelerated initial growth.[15] Other investigators have included additional semiempirical term or terms to the first-order interfacial oxidation kinetics or directly to the rate expression for dh/dt given by the linear-parabolic model.[16–19] From this brief account of some of the modifications to the linear-parabolic model, it is evident that the physical processes controlling the growth of the SiO_2 layer are not fully understood and a definitive model of this industrially important process has yet to be developed.

Blanc examined the possible causes in the failure of investigations based on t-h data to arrive at a definite model of the oxidation process.[20] He came to the conclusion that for many of the proposed models the available t-h data do not have the sensitivity to distinguish one model from another. He also reported that some physically distinct models can result in similar mathematical expressions relating h to t.

In kinetic investigations of chemical reactions, there is often an advantage in using model-independent rate of change in concentration $dc(t)/dt$ vs. concentration $c(t)$ data extracted from experimental time-concentration t-$c(t)$ data to identify the likely kinetic model and to determine the kinetic parameters in the model compared to the conventional method of using raw t-c data directly.[21] This observation is likely to remain valid for the silicon oxidation process in which model-independent plots of h-dh/dt derived from experimental t-h data are used instead of the raw t-h data. This approach may overcome some of the issues identified by Blanc.[20]

The rate of oxide growth dh/dt depends on the rate controlling step in the complex silicon oxidation process. If the slow rate of oxidation is the rate determining step then dh/dt is likely to be independent of h. Conversely, if the rate of supply of oxygen to the oxide-silicon interface is much slower than the oxidation rate then dh/dt would decrease with increasing h. Thus, the shape of the h-dh/dt plot provides a very vivid picture of changes in the rate controlling step of the diffusion-oxidation process. This is a further incentive for developing computation procedures for converting experimental t-h data into model-independent h-dh/dt plots.

It is essential that in the computation of a model-independent dh/dt the unavoidable measurement noise in the original t-h data is not amplified. Lubansky et al. demonstrated that Tikhonov regularization is a reliable procedure for converting experimental t-c data into dc/dt without amplifying measurement noise.[22] This investigation describes the extension of this procedure to t-h data in which a number of significant modifications are introduced. The procedure will then be used to convert a number of sets of experimental t-h data taken from the literature into model-independent h-dh/dt plots. The performance of a number of kinetic models of the of the diffusion-oxidation process in the literature are then used to describe these regulated h-dh/dt plots. Limitations of these models are discussed and modifications explored. The main aim here is to demonstrate the advantages of using model-independent h-dh/dt plots in model identification and parameter determination.

Integral Equation for the Derivatives

A typical set of oxide layer thickness measurements takes the form: $(\mathbf{t}^m, \mathbf{h}^m) = \{(t_1^m, h_1^m), (t_2^m, h_2^m), \ldots (t_i^m, h_i^m), \ldots (t_{N_D}^m, h_{N_D}^m)\}$. In some of the earlier datasets N_D is relatively small, between 10 and 30 points. With the development of new measurement techniques some experimenters reported data with N_D as large as 500–600 points covering a time span spreading over 3 to 4 decades. The superscript m is used here to distinguish the experimental data from the computed results which will carry the superscript c. The main computation task is to convert experimental $(\mathbf{t}^m, \mathbf{h}^m)$ measurements into instantaneous growth rate dh/dt without making any assumption regarding the underlying kinetics. Direct numerical differentiation, for example, by finite difference is unlikely to be satisfactory as this would amplify the unavoidable noise in the data. Tikhonov regularization as implemented by Lubansky et al. will be adapted to compute dh/dt.[22] A critical step of this method is the regularization parameter λ used to keep noise amplification under control. Detailed description of this method can be found in Engl et al.[23] and Lubansky et al.[22] A brief outline of the steps involved highlighting the modifications introduced is given later.

Tikhonov regularization computation used to obtain the second derivative d^2h/dt^2 from the measured data $(\mathbf{t}^m, \mathbf{h}^m)$ does not assume either the functional form of the data or that of the resulting second derivative. The computed second derivative is then integrated numerically to yield dh^c/dt which is in turn integrated to give h^c. All these results are independent of any assumed kinetic model. For convenience the computed first and second derivatives will be denoted by $r(t) = dh^c/dt$ and $f(t) = d^2h^c/dt^2$, respectively.

The second derivative is related to the computed thickness h^c by[24]

$$h^c(t) = \int_{t'=t_1}^{t} (t-t')\mathrm{f}(t')dt' + h_1 + tr_1 \tag{3}$$

This relationship is exact and the independent variable t varies from $t = t_1^m$ to $t = t_{N_D}^m$. In the computation scheme, Eq.

3 is treated as an integral equation of the first kind to be solved for $f(t)$ and h_1 and r_1—the estimated thickness and its first derivative at $t = t_1$, respectively. The time span of the experimental data under investigation is discretized into regularly spaced points $\mathbf{t} = [t_1 = t_1^m,\ t_2\ .,..\ t_j,\ \ldots..t_{N_K} = t_{N_D}^m]^T$ and the unknown function $f(t)$ at these points is represented by the unknown column vector $\mathbf{f} = [f_1,\ f_2,..\ f_j,\ldots\ f_{N_K}]^T$. N_K is the number of discretization points, typically $N_K >> N_D$. When $\mathbf{f}$ is substituted into Eq. 3 and the integral approximated by the Simpson's rule, this yield a linear algebraic equation for $h^c(t)$, at each of the measurement times $\mathbf{t}^m$, in terms of the unknowns $\mathbf{f}$ and r_1 and h_1. The computed values of r_1 and h_1, needed in subsequent computation, will be referred to as the computed initial rate and initial thickness, respectively.

The unknowns $[\mathbf{f},\ r_1,\ h_1]^T$ are required to minimize

$$(1)S_1 = \sum_{i=1}^{N_D} \left(h_i^m - h_i^c\right)^2$$
$$(2)S_2 = \sum_{j=2}^{N_K-1} \left(\frac{d^2f}{dt^2}\right)_{t_j}^2 \tag{4}$$

Condition (1) ensures that $h^c(t)$ approximates $h^m(t)$ closely and condition (2) ensures that $f(x)$ does not show spurious fluctuations. In Tikhonov regularization, instead of minimizing (1) and (2) separately, a linear combination $R = S_1 + \lambda S_2$ is minimized.[23] λ, the regularization parameter, balances these two requirements. A large λ favors the smoothness condition, whereas a small λ favors the accuracy condition.

For a given λ, minimizing R resulted in a set of linear algebraic equation for the unknowns $[\mathbf{f},\ h_1,\ r_1]^T$. As implemented by Lubansky et al., through a series of standard matrix operations, these unknowns are given explicitly by inversion of a matrix. It was found that, for some sets of time-thickness data, particularly those with less than 30–40 data points, the resulting dh/dt may not be monotonically decreasing. This is against the physical expectation of the growth of oxide layer. In such cases, an additional condition at each discretization point

$$\left(\frac{d^2f}{dt^2}\right)_{t_j} < 0 \ \text{for}\, j = 1 \ \text{to}\ N_K \tag{5}$$

has to be imposed. As a consequence, the unknowns $[\mathbf{f},\ h_1,\ r_1]^T$ can no longer be obtained by matrix manipulation. Instead numerical minimization of R now has to be performed. This is performed using well established numerical minimization procedure with imposed constraints. The procedure of Lubansky et al. also has to be modified when dealing with data that spread over 3–4 decades of oxidation time. For such large time span, a uniform discretization grid in time will lead to inadequate representation of data at small t. Here, the Tikhonov regularization computation is carried out with a uniform grid in the independent variable $s = \log\ t$ instead of t. The use of logarithmic variable is a standard practice when dealing with experimental data spreading over more than 4 decades.[25]

The choice of λ depends on N_D and N_K and on the noise level in the data.[23] In this investigation, the choice of λ will be guided by generalized cross validation (GCV) as described by Wahba.[26] In cases where GCV does not work, λ is selected so that the average deviation of $h^c(t)$ from $\mathbf{h}^m$ is comparable with the estimated error bars of the experimental data and such that the resulting derivatives do not show unrealistic fluctuations. This is referred to as the Morozov principle in some mathematical text.[23]

Oxide Layer Thickness Data and Growth Rate

Equations 3–5 will now be used to convert a number of sets of oxide layer thickness data into model-independent rate of growth. All the datasets analyzed are for dry thermal oxidation. Oxidation conditions and other experimental details associated with these data can be found in the original articles and will not be repeated here. All the Tikhonov regularization computations reported later, including the GCV curves, were performed after the thickness data and all unknowns have been converted into dimensionless form.

Data of Deal and Grove

Deal and Grove published several sets of $(\mathbf{t}^m,\ \mathbf{h}^m)$ data in 1965.[3] These have been analyzed extensively and formed the basis of a large number of progressively more complex modeling studies of the diffusion-oxidation process. The discrete points in Figure 1a are the data of these authors for the oxidation of silicon (111) at 700°C. These were reproduced by scanning and digitizing a large scale plot.[1]

When the matrix procedure of Lubansky et al. was applied to this set of $(\mathbf{t}^m,\ \mathbf{h}^m)$ data the resulting second derivative takes on positive values toward the end of the time span. This means that dh^c/dt is not monotonic decreasing and is, therefore, physically unacceptable. As aforementioned, the additional constraint Eq. 5 has to be imposed to ensure a monotonic decreasing rate. Numerical minimization, instead of matrix operations, was used to generate the set of unknowns $[\mathbf{f},\ h_1,\ r_1]^T$ that minimizes R. The resulting model-independent d^2h^c/dt^2 is shown in Figure 1b as a dark continuous curve. $N_K = 81$ for this result. The regularization parameter $\lambda = 10$ (dimensionless) is based on the minimum exhibited by the GCV curve in Figure 1c. It should be stressed that the numerical value of λ is not a physical property of the oxidation process as its value depends on the noise in the data as well as on N_D and N_K.

The monotonic first derivative dh^c/dt obtained by integrating d^2h^c/dt^2 numerically is shown in Figure 1d and the corresponding $h^c(t)$ obtained by similar computation from dh^c/dt is shown as a dark continuous curve in Figure 1a. In these computations, the rate and thickness $(r_1,\ h_1)$ generated by Tikhonov regularization were used as initial conditions. All these results are independent of any assumed model. The average deviation between $\mathbf{h}^c$ and $\mathbf{h}^m$ in Figure 1a is 4.59%. To reveal explicitly the variation of oxide growth rate with thickness, dh^c/dt is replotted against h^c in Figure 1e as a dark continuous curve. This curve clearly is also independent of any assumed model of the oxidation process.

Reisman et al. proposed a power-law expression to model the silicon oxidation process[27]

$$h^c = a_1(t + a_2)^{a_3} \tag{6}$$

where a_1, a_2, and a_3 are the parameters of the model. This yields a very simple expression for dh^c/dt

$$\frac{dh^c}{dt} = {a_1}^{1/a_3} a_3 \times h^{c(a_3-1)/a_3} = a_4 h^{c a_5} \tag{7}$$

a_4 and a_5 are directly related to a_1 and a_3 but for computation purpose it is more convenient to treat them as the new

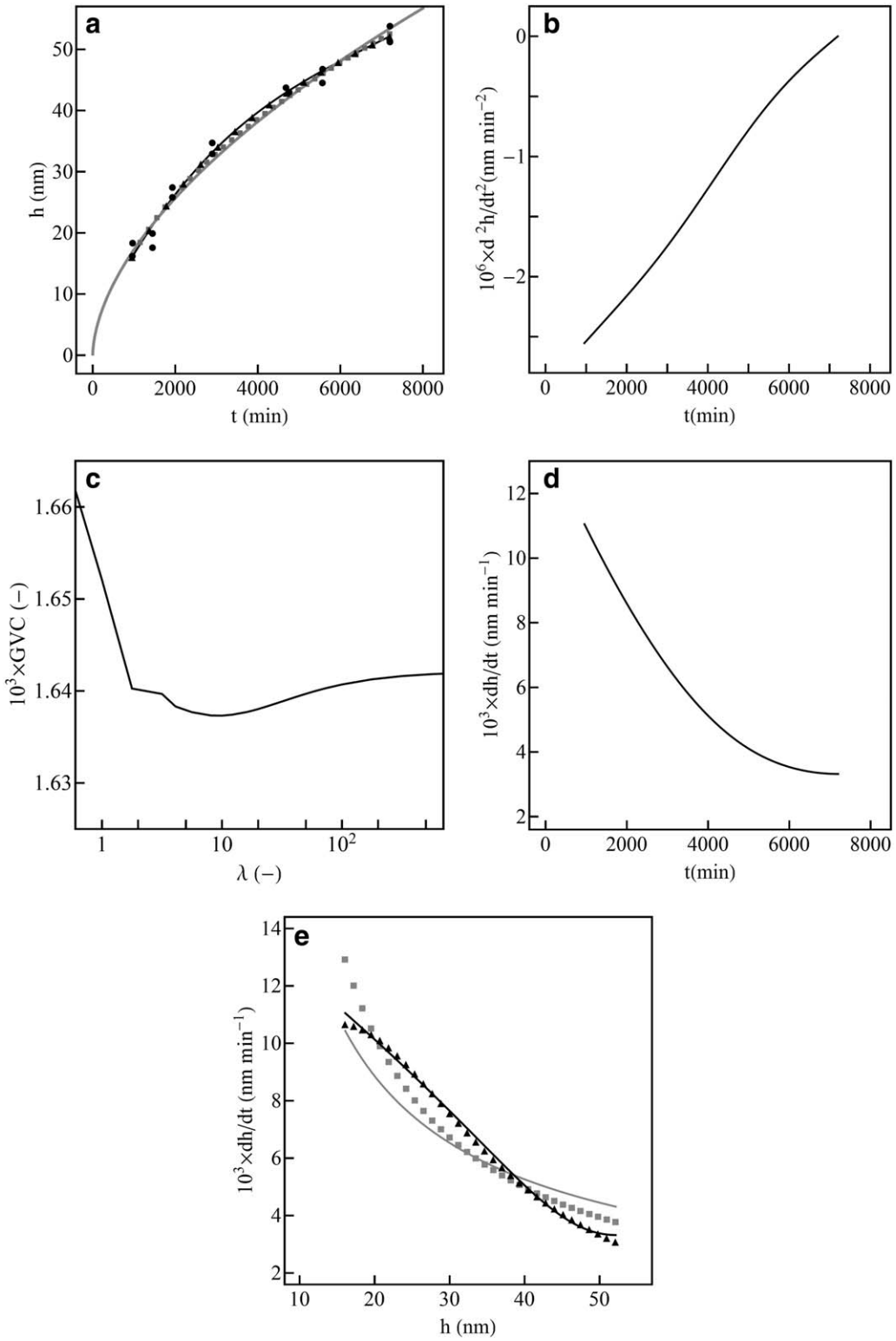

Figure 1. Deal and Grove Data.[1,3]

(a) Comparison of measured and computed thicknesses. Discrete points: experimental data; dark continuous curve: model-independent thickness from Tikhonov regularization; light continuous curve: thickness from the power-law model using Fair's parameters[1] Eq. 8; light squares: back-calculated from best-fit power-law rate expression Eq. 10; dark triangles: back calculated from rate expression Eq. 13 based on Blanc's model. (b) Temporal plot of the second derivative given by Tikhonov regularization. (c) Dimensionless GCV plot showing a minimum at $\lambda = 10$. (d) Temporal plot of the first derivative from Tikhonov regularization. (e) Spatial plots of the first derivative. Dark continuous curve: model-independent from Tikhonov regularization; light continuous curve: from power-law model Eq. 9 based on Fair's parameters[1]; light squares: best-fit power-law rate expression Eq. 10; dark triangles: best-fit rate expression Eq. 13 based on Blanc's model.

independent parameters. With Eq. 6, Fair reported the following expression[1]

$$h^c = 0.335(t + 0.40521)^{0.571} \tag{8}$$

for the thickness data in Figure 1a, where t is in minutes and h^c is in nm. For comparison, this is shown on the same plot as a light continuous curve. Following Fair, this curve is plotted over a time span larger than the available experimental data. Using the values of a_1 and a_3 reported by Fair to evaluate a_4 and a_5 in Eq. 7 led directly to

$$\frac{dh^C}{dt} = 0.084109 h^{c\,-0.751313} \tag{9}$$

This rate of growth, as required by physical expectation, is monotonic decreasing and is shown in Figure 1e as a light continuous curve. There is a very significant difference between this model-based rate curve and the model-independent rate curve.

Instead of computing the values of a_4 and a_5 from the values of a_1 and a_3 as reported by Fair they can also be determined directly by standard least-squares minimization between Eq. 7 and the model-independent curve in Figure 1e giving

$$\frac{dh^C}{dt} = 0.235505 h^{c\,-1.04461} \tag{10}$$

The units of t and h^c are the same as those in Eqs. 8 and 9. This expression is significantly different from Eq. 9 and is shown as a curve of light squares in Figure 1e. It is closer to but still significantly different from the model-independent dh/dt curve generated by Tikhonov regularization. However, when Eq. 10 is solved, as a first-order ordinary differential equation (ODE) for h^c with h_1 as the initial condition, the resulting model-based h^c is again in satisfactory agreement with the original data. This is shown as a curve of light squares in Figure 1a.

The two lighter curves in Figure 1e illustrate clearly that the rate expressions given by Eqs. 9 and 10 are unable to provide a satisfactory description of the growth rate vs. thickness result generated by Tikhonov regularization suggesting that more general rate expressions are called for. In his Variable Diffusivity Model, Blanc assumed that the diffusivity of O_2 changes through the oxide layer and arrived at the following expression[20]

$$t = c_0 + c_1 h^c + c_2 h^{c2} + c_3 \ln\left[1 + c_4 h^c\right] \tag{11}$$

that gives the oxidation time as a function of oxide thickness. In this investigation, the parameters c_0, c_1 and so forth will be treated as unknowns to be determined numerically by least-squares minimization. They can, however, be expressed in terms of the various assumed material properties in the model proposed by Blanc.[20] Blanc has shown that Eq. 11 is capable of describing time vs. thickness data with a large number of data points. As expected, it is also able to fit the small number of points in dataset in Figure 1a (not shown). The objective here is to test the ability of the rate expression derived from Eq. 11 in describing the h-dh/dt curve generated by Tikhonov regularization.

Upon differentiation Eq. 11 yields a rate expression that can be reduced to the general form that is convenient for the purpose at hand

$$\frac{dh^c}{dt} = \frac{1 + a \times h^c}{b \times h^{c2} + c \times h^c + d} \tag{12}$$

Here, the parameters a, b, c, and d can be expressed in terms of the parameters in Eq. 11 and hence with the original parameters in Blanc's model. Least-squares minimization of the fractional deviations of dh/dt from the Tikhonov regularization curve in Figure 1e led to

$$\frac{dh^c}{dt} = \frac{1 + 0.00240 \times h^c}{0.001907 h^{c2} - 0.005629 h^c + 0.138528} \tag{13}$$

where t and h^c are in minutes and nm, respectively. This expression is shown as a curve of dark triangles in Figure 1e. It is in much better agreement with the model-independent curve than the rate expressions Eqs. 9 and 10 even though differences still remain, particularly at the two ends of the data span. The back-calculated h^c based on this expression is shown as a curve of dark triangles in Figure 1a. It is, as expected, very close to the model-independent dark continuous curve and to the experimental data.

Data of Hopper et al.

Hopper et al. used *in situ* ellipsometry to measure the oxide layer thickness in dry thermal oxidation of silicon (100) and (111) over a range of temperature and O_2 pressure.[28] Steps were taken to remove the native SiO_2 layer before initiating the thermal oxidation process. In their ($\mathbf{t}^M$, $\mathbf{h}^M$) data, tabulated by Blanc, t^M ranged from less than 10 s to just over 47,000 s and the reported h^M ranged from 1 nm to just over 90 nm.[20] As with the data of Deal and Grove these data have been the subject of a number of modeling investigations.

The large time span of Hopper's data means that a uniform discretization grid in time would lead to inadequate representation of the data at small t. This is overcome by introducing the logarithmic variable $s = \log_{10} t$ as the independent variable and the ($\mathbf{t}^m$, $\mathbf{h}^m$) data are converted into ($\mathbf{s}^m$, $\mathbf{h}^m$) prior to performing Tikhonov regularization computation. The results are then converted back to the physical variables d^2h^c/dt^2, dh^c/dt, and h^c. It was found that the dh^c/dt curve given by the matrix procedure of Lubansky is monotonic without the necessity of imposing the additional condition in Eq. 5. This resulted in some savings in computing effort.

A typical set of Hopper's data (set 7 of Blanc's tabulation), for silicon (111) oxidized at 870°C and O_2 pressure of 1.56 atm, is shown as discrete points in Figure 2a. This set has been analyzed in detail by Blanc.[20] The number of data point $N_D = 91$. The d^2h^c/dt^2 generated by Tikhonov regularization is shown in Figure 2b. The horizontal axis, in this and in a number of subsequent plots, is displayed in the logarithmic scale to reveal the details of rapid initial growth. The corresponding model-independent dh^c/dt and h^c obtained by integration are shown in Figures 2c, a, respectively, as continuous curves. The number of discretization points in the logarithmic variable s is $N_K = 801$. The GCV curve for this set of data does not exhibit a minimum and, therefore, provide no guidance in the choice of the regularization parameter λ. In this case λ was based on the Morozov principle. In Figure 2a, the average deviation between the h^c and the experimental data is around 1.42% with the error being largest close to the start of the oxidation process. As in the previous example, dh^c/dt is plotted against h^c in Figure 2d. These plots show that dh^c/dt is indeed monotonic decreasing over the time and oxide thickness spans covered by the experimental data. In agreement with Blanc's observation, there does not appear to be a significant constant rate region as is usually assumed by analyses based on the Deal–Grove model.[20]

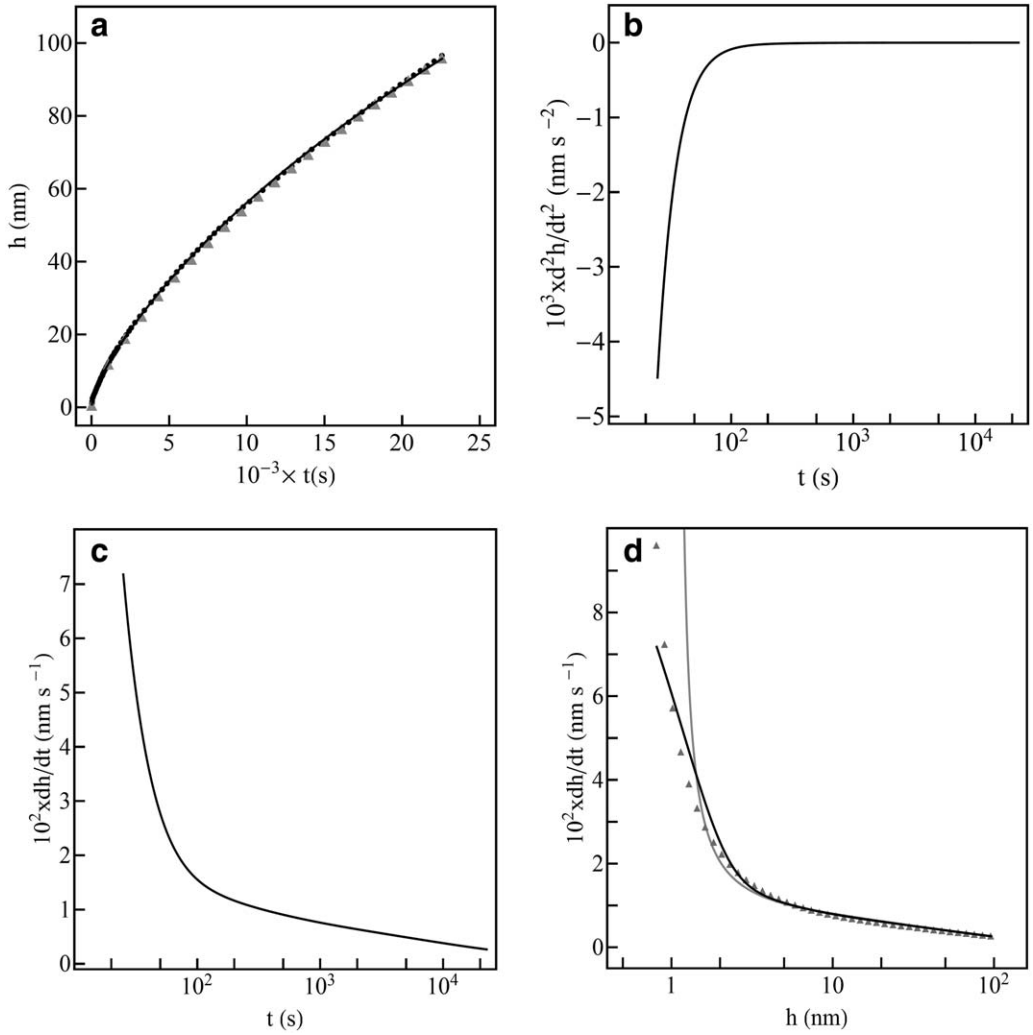

Figure 2. Data of Hopper et al.[28]

(a) Comparison of measured and computed thicknesses. Discrete points: experimental data; dark continuous curve: model-independent thickness from Tikhonov regularization and also best-fit Blanc's model[20]; light triangles: back-calculated from best-fit rate expression Eq. 17 based on Blanc's model. (b) Temporal plot of the second derivative given by Tikhonov regularization. (c) Temporal plot of the first derivative from Tikhonov regularization. (d) Spatial plots of the first derivative. Dark continuous curve: model-independent from Tikhonov regularization; light continuous curve from Blanc's model[20]; light triangles: best-fit rate expression rate expression Eq. 17 based on Blanc's model.

Adopting Blanc's variable-diffusivity model and minimizing the fractional deviations between Eq. 11 and the $\mathbf{t}^m$ of the data in Figure 2a led to

$$t = \frac{146.8693 + 111.0845\ h^c + 1.3968\ h^{c2} - 155.2522}{\ln(1 + 3.5056\ h^c)} \tag{14}$$

where t is in seconds and h^c is in nm. The curve given by Eq. 14 is essentially indistinguishable from the model-independent dark continuous curve in Figure 2a. For clarity, this new curve has been left out.

Differentiation of Eq. 14 followed by algebraic manipulation led to

$$\frac{dh^c}{dt} = \frac{1 + 3.50559\ h^c}{9.79334\ h^{c2} + 392.211\ h^c - 433.167} = \frac{0.01154}{h^c - 1.0755} + \frac{0.3464}{h^c + 41.125} \tag{15}$$

This gives the growth rate, in nm s^{-1} explicitly in terms of h^c in nm. It is plotted as a light continuous curve in Figure 2d. Over part of the thickness span, it is in satisfactory agreement with the model-independent dh/dt curve generated by Tikhonov regularization. However, because of the singularity at $h^c = 1.0755$ nm, Eq. 15 deviated greatly from the model-independent curve for small h^m. The light continuous curve has to be terminated before the singularity.

In their Parallel Oxidation Model, Han and Helms proposed that there are two oxidizing species responsible for the formation of SiO_2.[16] They arrived at a two-term rate expression of the general form

$$\frac{dh^c}{dt} = \frac{A_1}{h^c + B_1} + \frac{A_2}{h^c + B_2} \tag{16}$$

The parameters A_1, A_2, B_1, and B_2 in this expression are closely related to but not identical with the original parameters of Han and Helms. The two terms represent the separate contribution of the two oxidizing species to the overall rate. Blanc integrated the rate expression in Eq. 16 and through a series of algebraic manipulations arrived at an expression with exactly the same functional form as Eq. 11. Comparing

the integrated expression with Eq. 11 he was able to establish the relationships between the parameters in the two models in spite of the fact that they are based on entirely different physical assumptions.

However, it should be noted that, depending on the values of the numerical coefficients, it is not always possible to factorize the denominator in the general rate expression based on Blanc's model, Eq. 12, into two real factors as proposed by Han and Helms. This is the case with the rate expression in Eq. 13. And even when factorization is possible, as in Eq. 15, the resulting terms may not always give a positive growth rate individually as required by the model of Han and Helms. Thus, model-free h-dh/dt data have revealed that the two models are not strictly equivalent.

As in the previous example, instead of taking their values from Eq. 14, the four parameters in Eq. 12 were determined by least-squares minimization with the model-independent h^c-dh^c/dt curve in Figure 2d and the outcome is

$$\frac{dh^c}{dt}=\frac{1+0.27618\ h^c}{0.595655h^{c2}+44.344\ h^c-23.633}=\frac{0.02548}{h^c-0.5292}+\frac{0.4382}{74.9750+h^c} \tag{17}$$

where t and h^c are again in s and nm, respectively. This rate expression is displayed as a curve of light triangles in Figure 2d. This new model-based rate curve is in slightly improved agreement with the model-independent curve than Eq. 15. Deviation again becomes significant at low h^c as a consequence of the singularity which now lies outside of the span of the computed thickness h^c. Notwithstanding this significant deviation, when Eq. 17 is solved, as a first-order ODE for $h^c(t)$ with h_1 as the initial condition, the resulting h^c curve is in very good agreement with the original data of Hopper et al. This back-calculated $h^c(t)$ is shown as a curve of light triangles in Figure 2a. This suggests that a model capable of mimicking model-independent h-dh/dt data will generally yield a $h^c(t)$ curve that is in good agreement with ($\mathbf{t}^m$, $\mathbf{h}^m$) data.

Data of Enta et al.

Enta et al. used *in situ* ambient pressure x-ray photoelectron spectroscopy (APXPS) to investigate the growth of SiO_2.[29] This technique allowed them to measure oxide thickness in the range of 0–3 nm with an estimated precision of 0.1–0.2 nm. This thickness range is particularly relevant to the present day microelectronic industry. A typical set of the APXPS data for oxidation of Si(100) surface at a relatively low temperature of $T=300$°C and oxygen pressure $P=1$ Torr, based on the data files provided by Enta et al., is shown in Figure 3a where $N_D=583$. Figure 3b is an enlargement of the data points for small time t.

Using the logarithmic variable s, these data have been converted to the second derivative curve in Figure 3c with $N_K=801$. In the Tikhonov regularization computation, the first five data points with $h^m<0.15$ nm (approx.) have been ignored as these points have error bars too large to be handled by the regularization procedure and will result in negative growth rate or nonmonotonic decreasing rate for small t. See Figure 3b where the points left out are shown as dark squares. The choice of the regularization parameter for Figure 3c was guided by the minimum in the dimensionless GCV curve shown in Figure 3d. As in previous examples, the d^2h^c/dt^2 curve was integrated numerically to yield the dh^c/dt curve in Figure 3e and the h^c thickness vs. time curve in Figures 3a, b. The average deviation between the h^c curve and the $\mathbf{h}^m$ is 4.39% with the largest error at small h^c. Figure 3f shows the variation of dh^c/dt with h^c as a dark continuous curve.

Following previous examples, the general rate expression derived from Blanc's model in Eq. 12 was used to describe the rate vs. thickness relationship in Figure 3f. This has not been successful as the curve that minimized the fractional deviations does not show a monotonically decreasing relationship between h and dh/dt at small thickness. See curve of light squares in Figure 3f. This is physically unacceptable and can be taken as an indication that the growth rate there may require special treatment.

Recently Cui et al. examined the oxide growth data of Enta et al. and proposed a two-regime model to describe such data.[30] As the name implies, this model relies on two separate expressions to describe the variation of h^c with t

$$\begin{aligned} h^c&=d_1\sqrt{e^{2d_2t}-1}\ \text{ for } h^c\le h_{\text{trans}}\\ h^c&=\sqrt{d_3t+{d_4}^2}\ \text{ for } h^c>h_{\text{trans}}\end{aligned} \tag{18}$$

d_1, d_2, d_3, and d_4 are parameters that depend on the kinetic constants of the oxidation reactions and the diffusivities of the reacting species—either oxygen ion or molecular O_2 in the two regimes. These authors assumed that at some critical oxide thickness h_{trans}, there is an abrupt transition from the rapid growth observed in the thin oxide regime to the reduced growth rate in the thick oxide regime. h_{trans}, together with d_1, d_2, d_3, and d_4, forms the set of unknown parameters of the problem. As can be seen in Figure 3f, the dh^c/dt changes smoothly with h^c and does not exhibit an unambiguous h_{trans}.

The results of Tikhonov regularization computation, in particular the shape of the h^c-dh^c/dt curve, can be used to modify the two-regime model. Following previous examples, the two expressions in Eq. 18 are differentiated and rearranged to yield two rate expressions in terms of h^c.

$$\begin{aligned}\left(\frac{dh^c}{dt}\right)_I&=k_1h^c+\frac{k^2}{h^c}\\ \left(\frac{dh^c}{dt}\right)_{\text{II}}&=\frac{k_3}{h^c}\end{aligned} \tag{19}$$

The parameters k_1, k_2, and k_3 are related to the parameters in Eq. 18 but for computation purpose they are treated as new independent parameters. Instead of having the growth rate changing abruptly from $(dh^c/dt)_{\text{I}}$ to $(dh^c/dt)_{\text{II}}$ at some h_{trans}, it is assumed that there is a gradual transition from one to the other. Such a rate can be expressed as

$$\frac{dh^c}{dt}=G(h^c,k_4,k_5)\left(\frac{dh^c}{dt}\right)_I+[1-G(h^c,k_4,k_5)]\left(\frac{dh^c}{dt}\right)_{\text{II}} \tag{20}$$

The logistic function, $G(h^c,k_4,k_5)=1/(1+e^{k_4(h^c-k_5)})$ with two positive parameters k_4 and k_5, is used to link up $(dh^c/dt)_{\text{I}}$ and $(dh^c/dt)_{\text{II}}$.[31] For small h^c, the logistic function approaches 1, the composite rate approaches $(dh^c/dt)_{\text{I}}$. Conversely, as h^c becomes large, the logistic function approaches zero and the composite rate approaches $(dh^c/dt)_{\text{II}}$. k_4 controls the steepness of the transition and k_5 the general location of the transition, that is, a role similar to h_{trans}.

k_4 and k_5 together with k_1, k_2, and k_3 now form the set of parameters of the composite model. Although it

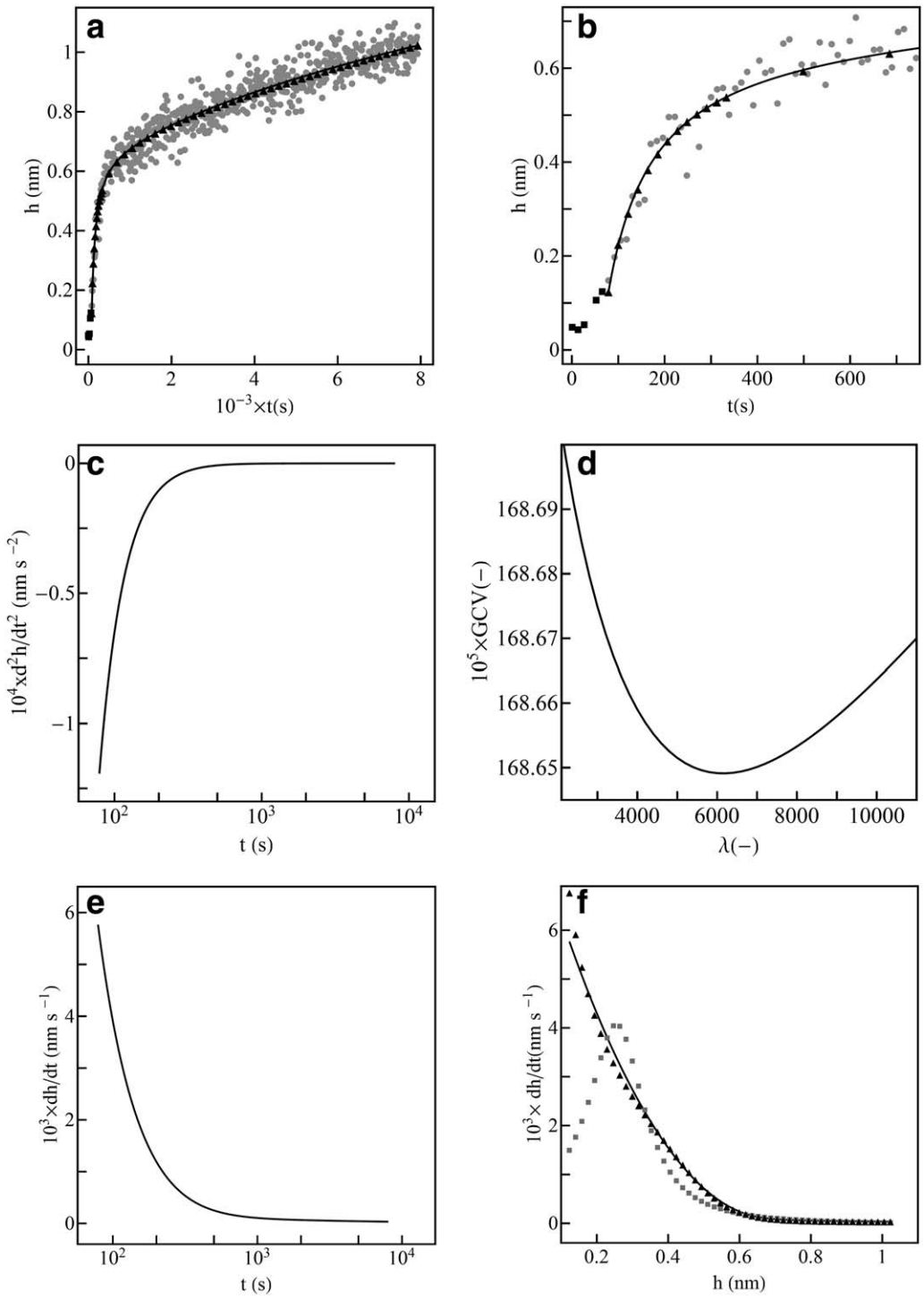

Figure 3. Data of Enta et al.[29]

(a) Comparison of measured and computed thicknesses. Light discrete points: experimental data; dark continuous curve: model-independent thickness from Tikhonov regularization; dark triangles: back calculated from best-fit composite rate expression Eq. 21. (b) Enlargement of (a) for small *t*. (c) Temporal plot of the second derivative given by Tikhonov regularization. (d) Dimensionless GCV plot showing minimum at $\lambda = 6200$. (e) Temporal plot of the first derivative from Tikhonov regularization. (f) Spatial plots of the first derivative. Dark continuous curve: model-independent from Tikhonov regularization; dark triangles: best-fit composite rate expression Eq. 21; light squares: best-fit rate expression Eq. 12 based on Blanc's model.

may be difficult to determine these parameters from the ($\mathbf{t}^m$, $\mathbf{h}^m$) data in Figure 3a, it is relatively simple to determine them by standard least-squares minimization of Eq. 20 and the model-independent h^c-dh^c/dt curve in Figure 3f. The outcome of the least-squares minimization is

$$\frac{dh^c}{dt} = G^*\left(\frac{0.084608}{hc} - 0.000098055\ hc\right) + (1-G^*)\left(\frac{0.0038111}{h^c}\right) \tag{21}$$

where $G^* = 1/(1+e^{1.50971(h_c-4.7566)})$ is the fitted logistic function. The t in Eq. 21 is in s and the h_c in Eq. 21 and G^* is

in nm. The composite rate expression is plotted as a curve of triangles in Figure 3f. The h^c curve back calculated from this expression is shown as a curve of triangles in Figures 3a, b. It is in satisfactory agreement with the model-independent curve. As in previous examples, notwithstanding the deviation in Figure 3f for small thicknesses, the back-calculated h^c is in very satisfactory agreement with the experimental ($\mathbf{t}^m$, $\mathbf{h}^m$) data and also with the model-independent curve in Figures 3a, b.

In the above computation, the logistic function was used to link the two separate rate expressions. Other sigmoid functions, of which the logistic function is a typical member, can equally be used to perform this function.[32] Justification and selection from this group of functions will require not only the expressions for different rate models but also physical details of the transition between them. This is likely to involve examination of the full diffusion-reaction equation instead of its pseudo steady-state simplification that is usually assumed. The full equation is a parabolic partial differential equation with a moving boundary.[33] Treatment of this equation is beyond the scope of the present investigation.[34]

Discussion

Tikhonov regularization originally developed to obtain reaction rates of chemical reactions from experimental concentration-time data has been extended to the diffusion-oxidation problem of silicon oxidation. With the imposition of the additional condition of negative second derivatives (if required), it has been successful in converting experimentally measured oxide layer thickness data into monotonic decreasing instantaneous rate of oxide growth. This success is attributed to the adjustable regularization parameter that maintains a fine balance between noise amplification inherent of numerical differentiation and the retention of as much of the physical information in the time-thickness data as possible.[22,35] This is a major improvement over unregularized numerical differentiation procedure such as the various finite difference schemes[36] or differentiation based on spline curves such the popular Savitzky–Golay method.[37]

From the way the Tikhonov regularization rates were obtained, it is clear that they are independent of any assumed functional relationship between time and thickness data and between rate and thickness. This is essential for the rates to be used in model testing. This would not be the case if the rates were obtained by differentiating curves that were fitted, either locally or globally, to match the experimental time-thickness data.

The rate of growth of the SiO_2 layer depends on the diffusion of the oxidizing species through the oxide layer and oxidation kinetics at the oxide-silicon interface. It is, therefore, not surprising that most theoretical investigations of the process resulted in expressions, such as Eqs. 7, 12, 19, and 20, relating the growth rate dh/dt to the oxide thickness h. Plotting model-independent growth rate dh/dt against h is thus the natural way of presenting experimental data for use in model testing and parameter determination.

Examination of the back-calculated h^c curves in Figures 1a and 2a shows that in each plot the rate models considered may be very different but all the model-based h^c curves are able to approximate the experimental data with similar degree of accuracy. This confirms Blanc's observation that available ($\mathbf{t}^m$, $\mathbf{h}^m$) data lack the sensitivity to distinguish one model from another.

Conversely, examination of the model-based dh^c/dt curves in Figures 1e, 2d, and 3f reveals a very different situation. h^c curves that are closed together do not necessarily have derivatives that are also close together. This is because differentiation amplifies the small differences between the h^c curves. When the model-based dh^c/dt curves are compared against their model-independent counterparts it becomes evident which rate model is able to provide, simultaneously, the most accurate description of the experimental h^m and the model-independent dh^c/dt. The h^c-dh^c/dt curves have succeeded where the ($\mathbf{t}^m$, $\mathbf{h}^m$) data have failed.

In all the cases investigated, the least-squares minimization of the difference between a rate model and the model-independent h^c-dh^c/dt curve converges smoothly to a well-defined set of model parameters with no significant numerical difficulty. The simplicity of this computation is particularly apparent in the case of the composite model in Eq. 20.

All these observations taken together back up the general assertion that matching model-independent h^c-dh^c/dt plots is a more stringent and yet simpler test of rate models of the silicon oxidation than raw ($\mathbf{t}^m$, $\mathbf{h}^m$).

Numerical experimentations showed that it is relatively simple to modify existing rate expressions for dh^c/dt, such as those investigated above, to improve their performance in describing the h^c-dh^c/dt curves generated by Tikhonov regularization. The ease with which such modifications can be implemented and tested more than justified the additional effort involved in the Tikhonov regularization computation. It is just as simple to use these regularized h^c-dh^c/dt curves to test completely new rate expressions of the oxidation-diffusion process. Such regularized rate curves make model testing so simple that it can easily degenerate into a mindless curve fitting exercise. Therefore, it is necessary to bear in mind that the ability of an assumed rate expression to describe a regularized curve h^c-dh^c/dt on its own does not make it an acceptable rate model. Unless the assumed functional form has a plausible physical explanation it has to be rejected. For example, a quick computer run will show that a fourth-order polynomial of the form

$$\frac{dh^c}{dt} = b_0 + b_1 h^c + b_2 h^{c2} + b_3 h^{c3} + b_4 h^{c4} \tag{22}$$

is, with different numerical parameters, able to describe the three Tikhonov rate curves in Figures 1e, 2d, and 3f with a degree of fidelity higher than that observed in this investigation. However, in the absence of physical justifications such an empirical rate expression cannot be accepted. In comparison, the rate expression derived from Blanc's model and the composite expression based on Cui's model, with their plausible physical interpretation, are models that can be considered for possible modification and further exploration with a larger number of datasets under different oxidation conditions.

A number of simple but crucial modifications of the method as implemented by Lubansky et al. had to be incorporated in the present investigation. The key one, where required, is that needed to ensure that growth is monotonic decreasing based on physical expectation. This was achieved by requiring the second derivative to be negative throughout the oxidation time span. This required the use of a numerical procedure instead of matrix operations to obtain the unknown second derivatives at the discretization points which increased slightly the computation time required. As

efficient numerical minimization procedures with constraints can be found in most scientific computing software, the additional programming effort is minimal.

Another essential modification is the introduction of the logarithmic variable $s = \log t$ to cope with the large range in time of the oxidation ($\mathbf{t}^m$ $\mathbf{h}^m$) data. The transformation between logarithmic variables and physical variables only requires a small number of simple numerical data management steps post Tikhonov regularization computation. These can be handled automatically within the computer program.

The regularization parameter is able to keep the amplification of noise in the data in check even when there is significant scatter in the thickness data. The method is able to obtain reliable model-independent growth rate of the oxide layer over a wide range of thickness, from a few nanometres to just over 95 nm. In this investigation where possible GCV was used to guide the choice of the regularization parameter λ. Where this is not possible the physically obvious Morozov Principle was used. In both cases, it is essential that the resulting derivatives should not exhibit excessive or unrealistic fluctuations. This clearly involves a certain degree of subjective judgement. This can be regarded as a means of incorporating the physical knowledge/understanding of the oxidation process into the computation process. Experience showed that provided the chosen λ is of the appropriate order of magnitude, small changes in λ do not greatly affect the resulting model-independent dh/dt. Fine tuning of this parameter is generally not required.[21]

Conclusions

Tikhonov regularization provides a reliable means of converting experimental time-thickness data into model-independent growth rates for the silicon oxidation process. Presented in the form of computed thickness vs. growth rate plots they provide an effective tool for model development and testing and for the numerical steps in parameter determination. This way of performing model identification is more sensitive than the traditional method of matching measured thickness against computed model-based thickness.

Acknowledgment

The authors thank Prof. Enta and Prof. Mun for making available to them the APXPS data files without which this investigation would not have been possible and also the anonymous reviewer whose comments led to significant improvements of the article.

Notation

A, B = Deal–Grove model parameters
A_1, A_2, B_1, B_2 = Han and Helms model parameters
G, G^* = logistic function
N_D = number of data point
N_K = number of uniformly spaced discretized point
R = linear combination of sum of squares
S_1, S_2 = sum of squares of difference
a, b, c, d = parameters in the rate expression based on Blanc's model
a_1, a_2, a_3 = parameters in power-law model for oxide thickness
a_4, a_5 = parameters in growth rate expression based on power-law model
b_0, b_1, b_2, b_3, b_4 = parameters in polynomial model for dh/dt
c = denotes concentration
c_0, c_1, c_2, c_3, c_4 = parameters in Blanc's model for oxide thickness
c = denotes computed value (superscript)
d_1, d_2, d_3, d_4 = parameters in Cui's model for oxide thickness
f = second derivative of $h(t)$
$\mathbf{f}$ = column vector of unknown discretized second derivative
h = thickness of oxidized layer
$\mathbf{h}$ = column vector of discretized thickness
h_{trans} = transition thickness in Eq. (18)
i, j = integer (subscript)
k_1, k_2, k_3 = parameters in expression for rate of growth based on Cui's model
k_4, k_5 = parameters in logistic function
m = denotes measured value (superscript)
r = first derivative of oxide thickness
s = $\log_{10} t$
$\mathbf{s}$ = uniformly discretized s
t = time
$\mathbf{t}$ = uniformly discretized time

Greek letters

λ = Tikhonov regularization parameter
τ = Deal–Grove model parameter

Literature Cited

1. Fair RB. Diffusion and oxidation of silicon. In: Hess DW, Jensen KF, editors. *Microelectronics Processing Chemical Engineering Aspects*. Washington, DC: ACS, 1989:265–323.
2. Middleman S, Hochberg AK. *Process Engineering Analysis in Semiconductor Device Fabrication*. Singapore: McGraw-Hill, 1993.
3. Deal BE, Grove AS. General relationship for the thermal oxidation of silicon. *J Appl Phys*. 1975;36:3770–3778.
4. Åkermark T. Molecular or atomic oxygen as the transported species in oxidation of silicon. *J Electrochem Soc*. 2000;147:1882–1887.
5. Bongiorno A, Pasquarello A. Oxygen diffusion through the disordered oxide network during silicon oxidation. *Phys Rev Lett*. 2002; 88:125901–125904.
6. Murali V, Murarka SP. Kinetics of ultrathin SiO2 growth. *J Appl Phys*. 1986;60:2106–2114.
7. Wang JM, Li Y, Li RW. An improved silicon-oxidation-kinetics and accurate analytic model of oxidation. *Solid State Electron*. 2003;47: 1699–1705.
8. de Almeida RMC, Gonçalves S, Baumvol IJR. Dynamics of thermal growth of silicon oxide films on Si. *Phys Rev B*. 2000;61:12992–12999.
9. Lu YZ, Cheng YC. A new model for the growth of silicon dioxide layers. *J Appl Phys*. 1984;56:1608–1612.
10. Fargeix A, Ghibaudo G, Kamarinos G. A revised analysis of dry oxidation of silicon. *J Appl Phys*. 1983;54:2878–2880.
11. Ajuria SA, Kenkare PU, Nghiem A, Mele TC. Kinetic analysis of silicon oxidations in the thin regime by incremental growth. *J Appl Phys*. 1994;76:4618–4624.
12. Gorantla S, Muthuvenkatraman S, Venkat R. A model for thermal growth of ultrathin silicon dioxide in O2 ambient: a rate equation approach. *IEEE Trans Electron Devices*. 1998;45:336–338.
13. Kageshima H, Shiraishi K. First-principles study of oxide growth on Si(100) surfaces and at SiO2/Si(100) interfaces. *Phys Rev Lett*. 1998; 81:5936–5939.
14. Watanabe T, Tatsumura K, Ohdoman I. New linear-parabolic rate equation for thermal oxidation of silicon. *Phys Rev Lett*. 2006;96: 196102–196104.
15. Singh SK, Schlup JR, Fan LT, Sur B. Modeling of thermal oxidation of silicon. *Ind Eng Chem Res*. 1988;27:1707–1714.
16. Han CJ, Helms CR. Parallel oxidation mechanism for Si oxidation in dry O2. *J Electrochem Soc*. 1987;134:1297–1302.
17. Kim K, Lee YH, An MH, Suh MS, Youn CJ, Lee K.B, Lee HJ. Growth law of silicon oxides by dry oxidation. *Semicond Sci Technol*. 1996;11:1059–1064.
18. Dimitriev S, Harrison HB. Modeling the growth of thin silicon oxide film on silicon. *J Appl Phys*. 1996;80:2467–2470.
19. Beck RB. Formation of ultrathin silicon oxides—modelling and technological constraints. *Mater Sci Semicond Process*. 2003;6:49–57.
20. Blanc J. The oxidation of silicon by dry oxygen can we distinguish between models? *Philos Mag B*. 1987;55:685–710.
21. Yeow YL, Wickramasinghe SR, Han B, Leong YK. A new method of processing the time-concentration data of reaction kinetics. *Chem Eng Sci*. 2003;58:3601–3610.

22. Lubansky AS, Yeow YL, Leong YK, Wickramasinghe SR, Han B. A general method of computing the derivative of experimental data. *AIChE J*. 2006;52:323–332.
23. Engl HW, Hanke M, Neubauer A. *Regularization of Inverse Problems*. Dordrecht: Kluwer Academic, 2000.
24. Burkill JC. *A First Course in Mathematical Analysis*. Cambridge: Cambridge University Press, 1978.
25. Yeow YL, Liow J-L, Leong Y-K. A general procedure for obtaining the evolving particle-size distribution of flocculating suspensions. *AIChE J*. 2012;58:3043–3053.
26. Wahba G. *Spline Models for Observational Data*. Philadelphia: SIAM, 1990.
27. Reisman A, Nicollian EH, Williams CK, Merz CJ. The modelling of silicon oxidation from 1 × 10−5 to 20 atmospheres. *J Electron Mater*. 1987;16:45–55.
28. Hooper MA, Clarke RA, Young L. Thermal oxidation of silicon in situ measurement of the growth gate using ellipsometry. *J Electrochem Soc*. 1975;122:1216–1222.
29. Enta Y, Mun BS, Rossi M, Ross PN Jr, Hussain Z. Real-time observation of the dry oxidation of Si(100) surface with ambient pressure x-ray photoelectron spectroscopy. *Appl Phys Lett*. 2008;93:0121101–0121103.
30. Cui H, Wang CX, Yang GW, Jiang D. Origin of unusual rapid oxidation process for ultrathin oxidation (<2 nm) of silicon. *Appl Phys Lett*. 2008;93:2031131–2031133.
31. Weisstein EW. Logistic equation. *MathWorld–A Wolfram Web Resource*. Available at http://mathworld.wolfram.com/LogisticEquation.html, accessed on 2014.
32. Weisstein EW. Sigmoid function. *MathWorld–A Wolfram Web Resource*. Available at http://mathworld.wolfram.com/SigmoidFunction.html, accessed on 2014.
33. Ghez R. *Diffusion Phenomena*. Dordrecht: Kluwer Academic, 2001.
34. Crank J. *Free and Moving Boundary Problems*. Oxford: Oxford University Press, 1984.
35. Yeow YL, Isac J, Khalid FA, Leong Y-K, Lubansky AS. A method for computing the partial derivatives of experimental data. *AIChE J*. 2010;56:3212–3224.
36. Khan IR, Ohba R. New finite difference formulas for numerical differentiation. *J Comput Appl Math*. 2000;126:269–276.
37. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem*. 1964;36:1627–1639.

Manuscript received May 8, 2013; and revision received Dec. 24, 2013.