

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

An investigation of electromigration induced void nucleation time statistics in short copper interconnects

V. M. Dwyer^{a)}

Department of Electronic and Electrical Engineering, University of Loughborough, Loughborough LE11 3TU, United Kingdom

(Received 29 October 2009; accepted 5 January 2010; published online 25 May 2010)

The stress evolution model (SEM) of Korhonen *et al.*, is used to calculate the void nucleation time in a large number of short interconnects (lengths up to 50 μm). Finite element calculations show that the effect of the nonlinearity in the SEM model is small, and that a mesh size of the order of the grain size is quite adequate to give accurate simulation results. Via failure is the only mode considered in the current calculations, however the gain in simulation time over other solution methods means that more complex situations, possibly including void dynamics, may be modeled in future in this way. Using normal mass-lumping methods the analysis is isomorphic to the voltage development on a random RC chain, so standard methods from very large scale integrated static timing analysis may be used to obtain dominant time constants at each mesh point. This allows the distribution of nucleation times to be obtained as a function of the distributions of line parameters. Under the assumption of a lognormal grain size distribution and a normal distribution of diffusion activation energies, the nucleation time distribution is shown to be close to lognormal. © 2010 American Institute of Physics. [doi:10.1063/1.3309744]

I. INTRODUCTION

The long time-to-failure under operational conditions means that electromigration (EM) failures can generally be reproduced in the laboratory only through accelerated life-tests.¹ Elevated values of the ambient temperature T and the current density j increase both the metal diffusivity and the EM driving force so that failures which would ordinarily only arise after years of use occur after perhaps hundreds of hours instead. Of course these tests are only useful, provided there is a robust mechanism both for deceleration from test conditions (of order 300 °C and 2 MA cm⁻²) back to operational conditions (of order 100 °C and 0.5 MA cm⁻²) and also for extrapolation, usually over several orders-of-magnitude, from the mean or median failure-times observed in the laboratory to the more useful reliability indicators, such as the time to 0.01% failures which is often required for reliability budgeting.² The utility of accelerated tests consequently relies on a good knowledge of the EM failure-time distribution and the dependence of that distribution on the acceleration parameters T and j . A second problem which arises in tests is that, even with such acceleration, it is very difficult to obtain a sufficiently large number of experimental results to give good statistics on any particular combination of metallization process and test structure. Nevertheless, it is fair to say that the majority of experimental results lie on a reasonably straight line, or at least are reasonably piecewise linear (possibly reflecting a number of different failure modes^{3,4}), when plotted on lognormal probability paper (for example, Refs. 5 and 6). As a result of this empirical fit, deceleration/extrapolation is generally performed based on lognormal statistics.⁷

In order to substantiate the claims of the lognormal dis-

tribution, it would be natural, given the difficulty in obtaining large quantities of experimental data, to turn to simulation studies. However the complex nature of the problem means that a full three-dimensional (3D) simulation, even of a single interconnect, may absorb many hours of CPU time. To illustrate the difficulty, a recently published 3D model simulation described results for a single interconnect of length around 2.5 μm , containing ten grains. Such simulations are vital in determining what is important in EM failure, but, being much shorter than the minimum length for failure, are unlikely to provide direct information about the failure-time distribution.⁸ Likewise such models cannot be extended to simulate lines of say 1000 μm length.

Despite such difficulties, there is much about EM which might be regarded as known. One such “known” is Black’s equation.^{9,10} Any strict lognormal distribution is, of course, completely determined by two parameters, the median failure-time t_{50} and the lognormal standard deviation σ_{TTF} (Ref. 11) and, provided that their dependence on acceleration parameters is known, the lognormal may be used to perform the deceleration/extrapolation process. Black’s equation seeks to describe the dependence of one of these, the median-time-to-failure t_{50} , on the acceleration parameters T and j , and is generally written in the form^{9,10}

$$t_{50} = \frac{A(T)}{j^n} \exp\left(\frac{E_A}{k_B T}\right). \quad (1)$$

Here E_A is an activation energy for the EM process and $A(T)$ appears to be, at most, a relatively weak function of temperature.^{7,12} Consequently E_A may be determined from a plot of $\log(t_{50})$ against T^{-1} . Black’s original empirical work gave a value of $n=2$ for the exponent of the applied current density j , while later experimental studies on aluminum interconnect structures produced values of n , for different

^{a)}Electronic mail: v.m.dwyer@lboro.ac.uk. Tel.: +44 1509 227027.

structures, generally lying between 1 and 3,¹²⁻¹⁴ but which were sometimes even as high as seven.¹⁵ The first theoretical studies,^{16,17} by contrast, suggested $n=1$, while later¹⁸⁻²² studies, which included back-diffusion of vacancies, suggested a value, in agreement with Black, of $n=2$. Empirically observed values of $n > 2$ have generally been accounted for as the result of Joule heating effects,¹³ while a spread of values between $n=1$ and 2 has been explained through a split of the failure-time t_f into a void nucleation contribution t_{nucl} (characterized by $n=2$) and a void growth contribution t_{grow} (characterized by $n=1$), i.e.,¹²

$$t_{50} = t_{\text{grow}} + t_{\text{nucl}} = \frac{A_{\text{grow}}(T)}{j} \exp\left(\frac{E_{\text{A}}^{\text{grow}}}{k_{\text{B}}T}\right) + \frac{A_{\text{nucl}}(T)}{j^2} \exp\left(\frac{E_{\text{A}}^{\text{nucl}}}{k_{\text{B}}T}\right). \quad (2)$$

In this way the combination of a particular line structure and metallization process which shows a slope of $n=1$ when $\log(t_{50})$ is plotted against $-\log(j)$, may be regarded as growth time dominated failure, while a slope of $n=2$ curve is nucleation dominated. These theoretical values may be derived in a simple manner from a number of related one-dimensional (1D) models including the stress evolution model (SEM) of Korhonen *et al.*,²⁰ which of such models, probably has the greatest acceptance at present. That a lognormal distribution with n close to one corresponds to a growth time dominated failure and a lognormal distribution with n close to two corresponds to a nucleation time dominated failure, implies that the distributions of both t_{nucl} and t_{grow} are both roughly lognormal also. For example, the system level EM simulator, SYSREL,²³ assumes that, for EM failure at vias between the first metal layer M1 and the second metal layer M2, $t_f = t_{\text{nucl}}$ (for a via-above structures²⁴) and $t_f = t_{\text{nucl}} + t_{\text{grow}}$ (for via-below structures²⁴), and that both are lognormally distributed.

The existence of a critical (Blech and Herring^{25,26}) length L_{B} , a length below which interconnect lines are immortal, is another known feature of EM which may also be explained by such 1D models, according to which the vacancy flux may be represented by²⁰

$$J = - \frac{D}{k_{\text{B}}T} \left(\frac{\partial \sigma}{\partial x} - G \right), \quad (3)$$

where $G = Z^* q \rho j / \Omega$. Here J is the vacancy flux, D the vacancy diffusivity, k_{B} is Boltzmann's constant, Ω is the atomic volume, Z^* the effective valance, q the electronic charge, and ρ the copper resistivity. Assuming a constant G at all points x for now, a zero flux condition $J = 0$, and hence a steady-state stress distribution $\sigma_{\text{ss}}(x)$, arises with

$$\sigma_{\text{ss}}(x) = \sigma_0 + \frac{Z^* q \rho j}{\Omega} x, \quad (4)$$

where σ_0 is the tensile stress at $x=0$. The maximum tensile stress in the line $\sigma_{\text{ss}}(x)$ is $\sigma_{\text{max}} = \sigma_0 + GL$, where L is the line length. Clearly if there is a critical stress σ_{cr} for void nucleation there will be a critical value of jL , obtained from setting $\sigma_{\text{max}} = \sigma_{\text{cr}}$. Interconnects with $jL < (jL)_{\text{crit}} = (\sigma_{\text{cr}} - \sigma_0) / (Z^* q \rho / \Omega)$ will then never nucleate a void and thus are essentially immortal. The existence of a Blech length is thus a natural consequence of these 1D models.

For a growing void, the free surface at the void boundary tends to keep the stress gradient small so that the flux into the void is dominated by the EM drift term, $J \approx DG / k_{\text{B}}T$. If a critical void volume Δ is required for failure then a growth time of $t_{\text{grow}} = \Delta k_{\text{B}}T / \Omega DG \sim 1/j$ is required. This yields a value of $n=1$ in Black's equation above. Most of the 1D models of EM failure approximate the flux by Eq. (3), the slight difference between them is in their treatment of the continuity equation. In the model of Korhonen *et al.*,²⁰

$$\frac{\partial \sigma}{\partial t} = - B \Omega \frac{\partial J}{\partial x}, \quad (5)$$

where B is an elastic modulus. Combining Eqs. (3) and (5) gives a parabolic equation similar to heat/diffusion type equations, but nonlinear due to an exponential dependence of the vacancy diffusion coefficient on the tensile stress σ ,

$$D(\sigma, x) = D_0(x) \exp\left(\frac{\sigma \Omega}{k_{\text{B}}T}\right). \quad (6)$$

If the interconnect is homogeneous (i.e., if D_0 independent of x) and if the model may be linearized, ($\sigma \Omega / k_{\text{B}}T \ll 1$), solutions are relatively straightforward and a value of $n=2$ is found for void nucleation. Essentially, in the linearized equation, σ/G obeys a diffusion-type equation with effective diffusion constant $\kappa = B \Omega D / k_{\text{B}}T$ and the tensile stress at the cathode-most end increases with a characteristic square-root time dependence, $\sigma(L, t) / G \sim \sqrt{\kappa t}$.²⁰ Thus if a critical stress σ_{cr} is required for nucleation then $t_{\text{nucl}} \approx \sigma_{\text{cr}}^2 / G^2 \kappa$, i.e., $t_{\text{nucl}} \sim j^{-2}$, corresponding to a value of $n=2$ in Black's equation. This analysis assumes a single mode failure at the cathode via. Values of n between 1 and 2 are thought to reflect interconnects in which both the nucleation time and the growth time are important contributors to the failure-time. As both t_{grow} and t_{nucl} behave as D^{-1} the parameter E_{A} in Black's equation may be simply related to the activation energy for diffusion.

While this seems convincing enough, there is something of a flaw in, or at least a major concern with this argument. The analysis we have gone through works only for homogeneous lines with position independent values of G and κ . It gives only a single value of t_{nucl} say ($\approx \sigma_{\text{cr}}^2 / G^2 \kappa$), or at least it gives one value for each set of values of σ_{cr} , G , and κ . How might this lead to a lognormal distribution? One option is to assume that the major variation comes from the diffusivity values at the copper/nitride-cap interface and that each interconnect may be modeled using an average value \bar{D} which could be used in the homogeneous-line expressions for t_{nucl} and t_{grow} . Perhaps \bar{D} is lognormally distributed between the lines which then will lead to a lognormal distribution of t_{nucl} and t_{grow} . However, if a grain's out-of-plane orientation determines its D value then, with each line containing a large number of grains, the central limit theorem (CLT) might be assumed to argue against \bar{D} being lognormal. In addition one might expect the intrasample variation in D values to be as great as the intersample variation.^{27,28}

Even more important than the failure-time distribution for a particular metallization process/test-structure is the failure-time distribution of a whole chip. This involves, at its simplest, connected tree structures of metal interconnect at a single metallization level, and a number of system level simulators have been introduced, the most recent of which SYSREL (Ref. 23) is also based around the SEM of Korhonen *et al.*²⁰ and the lognormal failure-time distribution. SYSREL makes the implicit assumption that the origin of the failure-time distribution comes from a lognormal distribution of average values \bar{D} , or at least a constant $D=D_k$ is applied to limb_k in the tree structure, and the limbs are subsequently treated as homogeneous.²³ Putting such concerns aside, it is clear that the model of Korhonen *et al.*²⁰ can be invoked to explain the origin of the Blech length, the nucleation/growth model which underpins the range of values for n seen in empirical fits to Black's equation and the system reliability modeling of programs such as SYSREL. It makes sense, as a result, to ask what such a model may say about the *distribution* of EM failure-times for a particular metallization process/test structure design, and in particular to ask to what extent the lognormal is a good fit to the failure times predicted by Korhonen *et al.*'s model.

With the picture of line failure described above, the failure-time of a test structure may then be determined by modeling the stress build-up in the interconnect, which results from the EM driving force. Specifically, on reaching some critical stress σ_{cr} a void nucleates. On nucleation, at t_{nucl} say, the stress is assumed to collapse to zero at the void's free surface, and the stress developed within the line, including the collapsed stress at the void, acts as the new initial condition for the growth stage of the process. Growth is assumed to have finished, and failure to have occurred, when the nucleated void (or the pre-existing void) grows to some critical size Δ . This critical void size will generally depend upon the geometry of the line.²⁹⁻³² For example a reservoir of copper material at the via will tend to increase the value of Δ for failure, while a reduced value will be required if the void nucleates within the via itself.³⁰ Of course if the line has a pre-existing void at the via, then $t_{nucl}=0$ and $t_f \cong t_{grow}$, while in so-called via-above (the void) structures, in which the void forms say in the M1 layer at an M1/M2 via, the void volume for failure is relatively small and $t_f \cong t_{nucl}$.²⁴ Even this is rather too simplistic as void dynamics is now recognized as a vital stage in failure; voids can form at points other than at the cathode via, they can depin from their nucleation site, drift in the interconnect, and combine with other existing voids. Nevertheless, first things first, there is an immediate question. In lines in which single mode failure is determined by the nucleation of a void at the cathode via and its subsequent growth to a critical failure size, what is the distribution of failure-times predicted by the Korhonen model? As indicated above, this could mean any of t_{nucl} , t_{grow} , or $t_{nucl}+t_{grow}$. We have considered the case of $t_{nucl}+t_{grow}$ for the typical via-below case previously³³ and turn to the distribution of t_{nucl} in the current work. Thus this work is most relevant to the via-above case,²⁴ or the via-below case²⁴ in which the void nucleates within the via itself,³⁰ i.e., cases in which Δ (and hence t_{grow}) is small.

The crystallographic orientation of grains on the cathode side of voids nucleated in line has been observed to show a distinctive trend, in that the majority have (111) out-of-plane orientations.^{27,28} In addition it is well known that lines with a strong (111) texture tend to have longer failure-times.³⁴ An obvious conclusion is that such grains have lower diffusivities than other orientations and act as semiblocking grains in much the same way that bamboo grains can block grain boundary diffusion in aluminum interconnect.³⁵ This suggests that the distribution of failure-times arises since identical test interconnects created on the same chip, and with the same process, possess different microstructures, each grain possessing its own diameter, orientation, diffusivity, and possibly other stochastic parameters such as adhesion strength and certain elastic properties. For this reason the critical stress σ_{cr} and critical void volume Δ may also differ for different interconnects, and even at different points on the same interconnect. To this extent it is a little misleading to suggest, as has been in the past, that the microstructure does not play a dominant role in the EM failure of Cu-based interconnects, as indeed, in order to describe failure statistics, it must play a fairly central role. What is generally meant by such statements is that the grain boundary network does not (yet) provide the dominant diffusion pathway, nevertheless the size and orientation of grains (currently at the copper/nitride-cap interface), and the resultant variety of interface diffusivity values, adhesion strengths, etc., that arise from the microstructure, must largely determine the distribution of interconnect lifetimes. This variety is intimately related to the microstructure of the lines, so that, in order to understand the influence of the microstructure on the failure-time distribution, the microstructure must be modeled in a realistic manner. A simple means of doing this for near-bamboo interconnects is given in Ref. 36. There, realistic microstructures may be developed, which possess the same statistical properties as patterned lines obtained using the two-dimensional grain growth simulator EMSim.³⁷ The grains created may be assigned parameters (size, effective diffusivity, effective valance, critical stress values, etc.) drawn from any distribution deemed appropriate. This may mean a lognormal grain size, a normal distribution of diffusion activation energies, fast and slow diffusivities corresponding to an activation energy with a normal mixture to model the effect of different crystallographic orientations and interface quality, and so on. The purpose of this paper is to investigate the void nucleation-time distribution in relatively short copper interconnects (up to around 50 μm in length) as a function of the distributions of the various material properties.

We seek first to solve the 1D stress evolution equation of Korhonen *et al.* using a simple finite element model (FEM). Because the critical stress σ_{cr} for void nucleation in copper appears to be relatively small, the size of the nonlinearity [$\sim \exp(\sigma_{cr}\Omega/kBT)$] is also small. As a result one might expect that linearizing the model is legitimate. However, there is a certain amount of nonlinear feedback in the vacancy dynamics, which means that this linearization still needs to be justified. Next, we investigate the effects of the FEM mesh-size on the results and consider some standard approximations used in FEM such as the lumping of the (equivalent

of the) mass matrix to simplify the problem. As a result nucleation time distributions are obtained under a variety of parameter assumptions. Finally, we consider an analytical approach.

Where this analysis differs from previous studies (e.g., Refs 18–23) of the SEM equation is that those studies generally assume that the lines are homogeneous, i.e., that the dependence, on the position x along the line, of the diffusivity, the adhesion strength,^{38,39} the bulk modulus, etc., are ignored. This latter simplification allows the “knowns” discussed above to be addressed but removes the possibility of modeling the failure-time distribution, since the details of the microstructure, in terms of a distribution of grain sizes and a distribution of single-grain diffusivity values, etc., necessarily requires the study of an inhomogeneous line. In particular, it is the distribution of nucleation times which is sought here.

II. GENERAL NONLINEAR, 1D FEM SOLUTION

In this section a FEM solution to the 1D, SEM equation is sought. As parameters such as the effective diffusivity κ , adhesion strength, etc., may vary with crystallographic orientation, such parameters will vary on the scale of the grain diameter. For maximum generality the effective valence is also allowed to be a stochastic variable so that, for a given interconnect, a value Z_k^* defines the effective valence for grain k , $k=1, 2, \dots, N_{\text{grains}}$. Similarly $\kappa_k (=B_k D_{0k} \Omega / k_B T)$ is the effective diffusivity of grain k which depends on the local bulk modulus B_k (and possibly adhesion strength^{38,39}) and the grain orientation (here D_{0k} is the zero-stress nitride/copper-interface diffusivity for grain k). Likewise, the diameter of grain k is labeled d_k . No particular boundary or initial conditions are specified at this time. For now all we assume is that there are no “boundary conditions” within the region $x \in [0, L]$ except for those at $x=0$ (the anode) and at $x=L$ (the cathode). Consequently to model void dynamics or an interconnect containing several voids, whose free surfaces pin the stress to zero at a number of interior points, it will be necessary to solve the problem in several parts and piece them together as necessary.

Combining Eqs. (3), (5), and (6), the stress evolution equation describes the hydrostatic tensile stress at (x, t) according to the nonlinear, parabolic partial differential equation²⁰

$$\frac{\partial \sigma}{\partial t} - B\Omega \frac{\partial}{\partial x} \left\{ \frac{D_0(x)}{kT} \exp\left(\frac{\sigma\Omega}{kT}\right) \left[\frac{\partial \sigma}{\partial x} - G(x) \right] \right\} = 0. \quad (7)$$

Introducing the dimensionless variable $\Sigma(x, t) = \sigma(x, t)\Omega / k_B T$, and the variables $\alpha(x) = \Omega G(x) / k_B T$ and $\kappa(x) = D_0(x)B\Omega / k_B T$, Eq. (7) may be written as

$$\frac{\partial \Sigma}{\partial t} - \frac{\partial}{\partial x} \left\{ \kappa(x) \exp(\Sigma) \left[\frac{\partial \Sigma}{\partial x} - \alpha(x) \right] \right\} = 0. \quad (8)$$

To simplify the analysis we keep the elastic modulus B constant for now, however, the inclusion of an inhomogeneous elastic modulus $B(x)$ is straightforward. Using standard Galerkin ideas,⁴⁰ we now approximate the normalized stress $\Sigma(x, t)$ by a sum of basis functions $\theta_k(x)$. Thus

$$\Sigma(x, t) \approx \hat{\Sigma}(x, t) \equiv \sum_{k=0}^N c_k(t) \theta_k(x), \quad (9)$$

where N is related to the size of the FE mesh. The approximation $\hat{\Sigma}(x, t)$ in Eq. (9) will not normally solve the problem exactly so that the right-hand side of Eq. (8), with $\Sigma = \hat{\Sigma}(x, t)$, will not normally be identically zero, and as a result there will necessarily be some residual error, denoted $R[\hat{\Sigma}] = R(x, t)$. As common with such methods, the coefficients $c_k(t)$ are calculated to make the residual orthogonal to $\hat{\Sigma}(x, t)$. Usually this is done by making $R[\hat{\Sigma}]$ orthogonal to each of the $\theta_k(x)$ in turn.⁴⁰ These orthogonality conditions are, for all k ,

$$\int_0^L \theta_k(x) R[\hat{\Sigma}] dx \equiv \int_0^L \theta_k(x) \left[\frac{\partial \hat{\Sigma}}{\partial t} - \frac{\partial}{\partial x} \left\{ \kappa(x) \exp(\hat{\Sigma}) \times \left[\frac{\partial \hat{\Sigma}}{\partial x} - \alpha(x) \right] \right\} \right] dx = 0, \quad (10)$$

Rearranging and integrating, Eq. (10) yields the (weak) integral formulation⁴⁰

$$\begin{aligned} & \frac{d}{dt} \int_0^L \theta_k(x) \hat{\Sigma}(x, t) dx \\ &= \theta_k(x) \kappa(x) \exp(\hat{\Sigma}) \left(\frac{\partial \hat{\Sigma}}{\partial x} - \alpha(x) \right) \Big|_0^L \\ & - \int_0^L \kappa(x) \exp(\hat{\Sigma}) \left(\frac{\partial \hat{\Sigma}}{\partial x} - \alpha(x) \right) \frac{\partial \theta_k(x)}{\partial x} dx, \end{aligned} \quad (11)$$

and it is this equation which we shall solve on the FE mesh.

We define the mesh \mathcal{N} as $N+1$ points x_k in the interval $[0, L]$, ordered such that $x_j < x_k$ if $j < k$. It is natural to choose $x_0=0$ and $x_N=L$, however the other points in \mathcal{N} may be chosen arbitrarily. This arbitrariness is one of the strengths of the finite element method as it allows more mesh points to be chosen where a solution function is changing rapidly. Here we choose them to define a uniform mesh \mathcal{M} on $[0, L]$ of step size ε , plus the set of grain boundary points \mathcal{G} , so that $\mathcal{N} = \mathcal{M} \cup \mathcal{G}$. Any resulting points in the ordered union \mathcal{N} which are too close (perhaps closer than $\varepsilon/2$) may be considered unnecessary and may be removed. This is assumed to be done in a manner which always retains the grain boundary points in \mathcal{G} over the regular mesh points in \mathcal{M} , and in any case never removes points in \mathcal{G} . In this way no grain boundary lies between any two adjacent points in the final mesh \mathcal{N} , so that the vacancy diffusivity $\kappa(x)$ will be constant κ_k over any interval of the form (x_{k-1}, x_k) , for any $x_{k-1}, x_k \in \mathcal{N}$. Similarly $G(x)$ is a constant G_k on the same interval. In addition $\varepsilon_k = x_k - x_{k-1}$ will, roughly speaking, be in the range $[\varepsilon/2, 2\varepsilon]$, although it is possible that the randomly generated grain structure will produce points in \mathcal{G} which are closer than $\varepsilon/2$. After construction of the mesh \mathcal{N} , we suppose there are $N+1$ points where roughly $N \sim L / \min(\varepsilon, d_{50})$.

The test functions $\theta_k(x)$ are chosen to be the well known hat (or chapeau) functions⁴⁰

$$\theta_k(x) = \begin{cases} \frac{x - x_{k-1}}{x_k - x_{k-1}} & \text{if } x_{k-1} < x \leq x_k \\ \frac{x_{k+1} - x}{x_{k+1} - x_k} & \text{if } x_k < x \leq x_{k+1} \\ 0 & \text{otherwise} \end{cases}, \quad (12)$$

except at the ends (x close to $x_0=0$ and x close to $x_N=L$) where

$$\theta_0(x) = \begin{cases} \frac{x_1 - x}{x_1} & \text{if } x \in [0, x_1] \\ 0 & \text{otherwise} \end{cases} \theta_N(x) = \begin{cases} \frac{x - x_{N-1}}{L - x_{N-1}} & \text{if } x \in [x_{N-1}, L] \\ 0 & \text{otherwise} \end{cases}. \quad (13)$$

Note that $\theta_k(x_m) = \delta_{km}$ (the Kronecker delta), and that, in terms of the basis functions given in Eqs. (12) and (13), Eq.

(9) implies that in the interval $x_{k-1} \leq x \leq x_k$, $\hat{\Sigma}(x, t)$ is given by

$$\hat{\Sigma}(x, t) = c_{k-1}(t)\theta_{k-1}(x) + c_k(t)\theta_k(x), \quad (14)$$

with $\hat{\Sigma}(x_k, t) = c_k(t)$. From Eq. (14), Eq. (11) may then be reduced to a series of equations for the time derivatives of $c_k(t)$, $k=0, 1, \dots, N$, the coefficients of the basis functions $\theta_k(t)$. These equations can be combined, as usual, into a single matrix equation, thus

$$\mathbf{M} \frac{d}{dt} \mathbf{c}(t) = -\mathbf{K} \exp \mathbf{c}(t) + \begin{bmatrix} 0 \\ \mathbf{KAF} \\ 0 \end{bmatrix} - \begin{bmatrix} \mathbf{KAF} \\ 0 \end{bmatrix} + \begin{bmatrix} J(0, t) \\ 0 \\ \vdots \\ 0 \\ -J(L, t) \end{bmatrix}, \quad (15)$$

where the “mass” and “stiffness” matrices \mathbf{M} and \mathbf{K} are just

$$\mathbf{M} = \frac{1}{6} \begin{bmatrix} 2\varepsilon_1 & \varepsilon_1 & 0 & \cdots & 0 & 0 \\ \varepsilon_1 & 2(\varepsilon_1 + \varepsilon_2) & \varepsilon_2 & 0 & \cdots & 0 \\ 0 & \varepsilon_2 & 2(\varepsilon_2 + \varepsilon_3) & \varepsilon_3 & 0 & \vdots \\ \vdots & 0 & \varepsilon_3 & \ddots & \varepsilon_{N-1} & 0 \\ \vdots & \vdots & \vdots & \varepsilon_{N-1} & 2(\varepsilon_{N-1} + \varepsilon_N) & \varepsilon_N \\ 0 & \cdots & 0 & 0 & \varepsilon_N & \varepsilon_N \end{bmatrix} \quad (16)$$

and

$$\mathbf{K} = \begin{bmatrix} \kappa_1/\varepsilon_1 & -\kappa_1/\varepsilon_1 & 0 & \cdots & 0 & 0 \\ -\kappa_1/\varepsilon_1 & \kappa_1/\varepsilon_1 + \kappa_2/\varepsilon_2 & -\kappa_2/\varepsilon_2 & 0 & \cdots & 0 \\ 0 & -\kappa_2/\varepsilon_2 & \kappa_2/\varepsilon_2 + \kappa_3/\varepsilon_3 & -\kappa_3/\varepsilon_3 & 0 & \vdots \\ \vdots & 0 & -\kappa_3/\varepsilon_3 & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & -\kappa_{N-1}/\varepsilon_{N-1} & \kappa_{N-1}/\varepsilon_{N-1} + \kappa_N/\varepsilon_N & -\kappa_N/\varepsilon_N \\ 0 & \cdots & 0 & 0 & -\kappa_N/\varepsilon_N & \kappa_N/\varepsilon_N \end{bmatrix}. \quad (17)$$

The values here are obtained by generating grain boundary diameters d_k , $k=1, 2, \dots, N_{\text{grains}}$, from (say) a lognormal size distribution, and then introducing regular mesh points of spacing ε within the resulting grains. Thus for example if a grain of diameter $d_1=0.515 \mu\text{m}$ is constructed between $x=0$ and $0.515 \mu\text{m}$, with a normalized diffusivity of 1.4 a.u. (arbitrary units), on a mesh of size $\varepsilon=0.1 \mu\text{m}$, we will have $\varepsilon_1=\varepsilon_2=\varepsilon_3=\varepsilon_4=0.1 \mu\text{m}$, and $\varepsilon_5=0.115 \mu\text{m}$ (as the mesh point at $0.5 \mu\text{m}$ will have been removed) and $\kappa_1=\kappa_2=\kappa_3=\kappa_4=\kappa_5=1.4$ a.u. The remaining vectors in Eq. (15): \mathbf{c} , $\exp(\mathbf{c})$, and \mathbf{KAF} are

$$\mathbf{c}(t) = [c_0(t), c_1(t), \dots, c_N(t)]^T,$$

$$\exp \mathbf{c}(t) = \{\exp[c_0(t)], \exp[c_1(t)], \dots, \exp[c_N(t)]\}^T,$$

$$\mathbf{KAF} = [\kappa_1 \alpha_1 F_{1,0}, \kappa_2 \alpha_2 F_{2,1}, \dots, \kappa_N \alpha_N F_{N,N-1}]^T, \quad (18)$$

where

$$F_{k,k-1}(t) = \frac{\exp[c_k(t)] - \exp[c_{k-1}(t)]}{c_k(t) - c_{k-1}(t)}, \quad (19)$$

For the relatively small systems (L less than around $50 \mu\text{m}$) under consideration here, the nonlinear SEM equation, Eq. (15), may be solved within MATLAB using the ODE solver ODE45 which uses an explicit fifth-order Runge–Kutta method, with fourth-order error control⁴¹ (alternatively, ode15s if the Eq. (15) shows signs of stiffness).

As an example, let us consider a pad-stud line, so that $\Sigma(x=0, t)=0$ and $J(x=L, t)=0$. Since Eq. (15) requires $J(x=0, t)$ to be known rather than $\hat{\Sigma}(x=0, t)$, in such cases as the

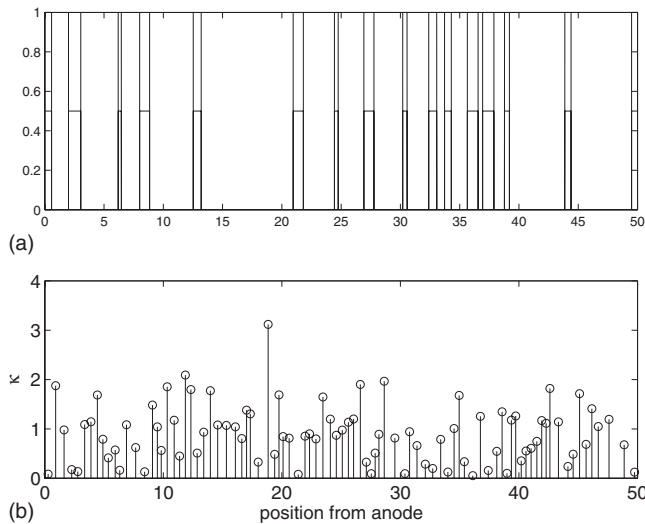


FIG. 1. (a) Fast and slow grains generated randomly. Fast grains are indicated with a line through the middle. (b) Diffusivities values generated from a normal mixture (fast and slow) of diffusion activation energies.

present one, the equation for $\Sigma(x=0, t)$ [or $c_0(t)$] is removed from the equation set and $\Sigma(x=0, t)$ is set equal to zero in the remaining equations. This simply corresponds to removing the first row and column of \mathbf{M} and \mathbf{K} , removing the first elements of each of the vectors in Eq. (15); and setting $c_0(t)=0$ in what remains. It is easy, although a little long-winded, to show that, in this case, Eq. (15) possesses the correct steady state solution of

$$c_k(t \rightarrow \infty) \equiv (c_\infty)_k = \sum_{j=1}^k \alpha_j \varepsilon_j. \quad (20)$$

For our example we choose $L=50 \mu\text{m}$ and the value of d_{50} , the median grain size, to be $0.5 \mu\text{m}$. The mesh size parameter ε is taken to be $0.1 \mu\text{m}$, so that a typical line has around 100 grains and 500 mesh points, and a typical grain contains around 5 mesh points. Grain diameters are chosen from a lognormal grain size distribution with lognormal standard deviation of 0.36.³⁸ In this example the activation energy which determines the effective diffusivity is drawn from a normal mixture. Grains are slow with probability $p=0.2$ and fast otherwise. In this way the values of κ_{eff} (fast grains) and κ_{eff} (slow grains) are lognormal. For now we arbitrarily choose a value of $\kappa_{50\text{fast}}/\kappa_{50\text{slow}}=8$ with lognormal standard deviations of 0.9 for both fast and slow grains. Microstructures are created using the technique in Ref. 36, and the grain sizes d_k and effective diffusivities κ_k are indicated in Fig. 1. Figure 1(a) indicates the position of the slow grains as blocks with a bar through the middle (the large regions correspond to several adjacent fast grains). The diffusivity values of each grain are shown in arbitrary units in Fig. 1(b), the interconnect is made of 89 grains in this case. Once the mesh was created there were a total of 543 mesh points so that an average \bar{d} (rather than d_{50}) sized grain contains around six mesh points.

Equation (15) is solved for this case and Fig. 2 (solid curve) shows the time dependence of the stress $\Sigma(x, t)$ at a variety of points (note the solution is obtained at all 543

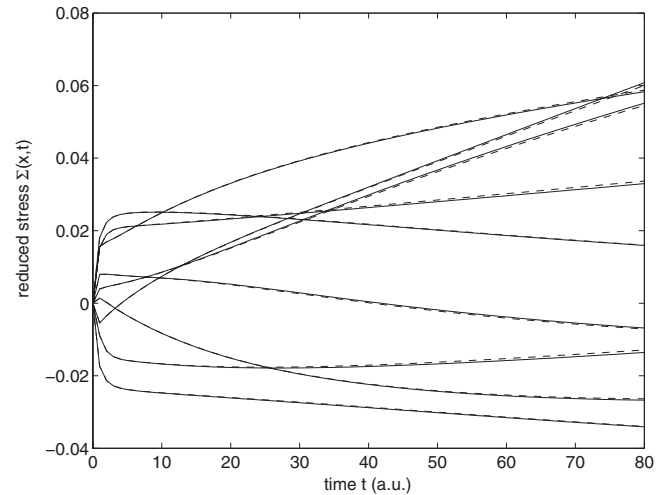


FIG. 2. Finite Element solution to the nonlinear and linear SEM equations. The nonlinear solution [solid curve, Eq. (15)] using the MATLAB routine ODE45 and the linear solution [dashed curve, Eq. (15) with Eq. (21)].

mesh points for this interconnect but only a few are shown here). Calculations using a smaller mesh parameter are possible but do become quickly intractable unless one also reduces L from $50 \mu\text{m}$. It is perfectly possible to simulate a $20 \mu\text{m}$ line with 20 mesh points per typical grain and 800 mesh points overall. However as we shall see a linearization of the problem produces an almost identical answer, and much greater progress may be made if such an approximation can be legitimized.

There are several linearization schemes possible for this system. The first is to approximate $\exp[c_k(t)]$ by $[1+c_k(t)]$ and to set

$$F_{k,k-1}(t) = \frac{\exp[c_k(t)] - \exp[c_{k-1}(t)]}{c_k(t) - c_{k-1}(t)} \approx 1, \quad (21)$$

in Eq. (15). If this is done, the solution to the same problem (Fig. 1.) is shown as the dashed curve in Fig. 2. The difference between linear and nonlinear solutions is never greater than 1% of the final voiding stress. In many ways this is to be expected as $\exp(\Omega\sigma_{\text{cr}}/kT)$ is close to unity, nevertheless it is important to demonstrate the fact. A second possible linearization is to approximate

$$F_{k,k-1}(t) = \frac{\exp[c_k(t)] - \exp[c_{k-1}(t)]}{c_k(t) - c_{k-1}(t)} \approx 1 + \frac{c_k(t) + c_{k-1}(t)}{2}, \quad (22)$$

as this is of similar order to the approximation of $\exp[c_k(t)]$ by $1+c_k(t)$. Numerically there is no observable difference between the two. From an analytical perspective the difference is that with Eq. (22), the \mathbf{K} matrix, although still tridiagonal, loses its symmetry as factors $(1 \pm \alpha_k \varepsilon_k/2)$ are introduced above and below the main diagonal. As the $\alpha_k \varepsilon_k$ values are small the solutions are almost identical again. The linear and nonlinear solutions have been compared on a mesh of $\varepsilon=0.1 \mu\text{m}$ for a large number of $50 \mu\text{m}$ microstructures and the results all show similar agreement. Equation (21) is adopted here, for use later in analytical development.

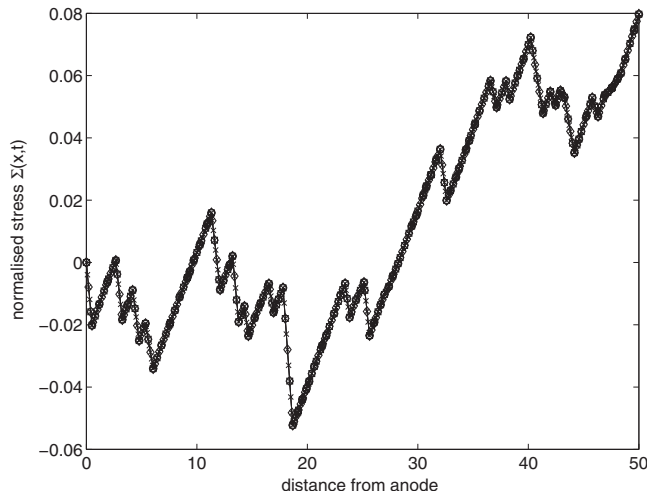


FIG. 3. Comparison of FEM solutions to the linearized equation with a variety of mesh sizes ranging from around $d_{50}/5$ to d_{50} .

III. LINEARIZED SOLUTIONS AND THE CUMULATIVE DISTRIBUTION FUNCTION (CDF)

As demonstrated above, solutions to the linearized equation are just as accurate as solutions to the original nonlinear one. Certainly the error made through assuming that the stress evolution equation of Korhonen *et al.*, (or any other analytical equation) is appropriate to EM will probably more than mask any error which occurs in neglecting the nonlinearity. As usual, though, linear problems submit to easier analysis and often display greater numerical stability. In particular we may now obtain solutions using matrix methods. Equation (15) becomes

$$\mathbf{M} \frac{d}{dt} \mathbf{c}(t) = -\mathbf{K}(\mathbf{c}(t) - \mathbf{c}_\infty), \quad (23)$$

so that

$$\mathbf{c}(t) = [\mathbf{I} - \exp(-\mathbf{M}^{-1}\mathbf{K}t)]\mathbf{c}_\infty \quad (24)$$

Using standard MATLAB matrix methods in Eq. (24), gives identical solutions to the ODE solver ODE45 but is naturally much quicker. This simplification is important as it also allows us to investigate, in a much easier manner, the number of mesh points (i.e., the value of ε) which is required to give an accurate solution to the problem. The same problem, shown in Fig. 1, is solved using Eq. (24) for values of $\varepsilon = 0.1 \mu\text{m}$, $0.2 \mu\text{m}$, $0.4 \mu\text{m}$, and ε large (which corresponds to $\mathcal{N}=\mathcal{G}$). The results are shown in Fig. 3, an arbitrary section of which is shown in Fig. 4. Clearly there is again no difference between the plots. This arises as the stress distributions are virtually linear across each grain. Repeated trials on different microstructures show the same behavior and as a result the mesh need only be defined at grain boundaries, thus it is sufficient to take $\mathcal{N}=\mathcal{G}$. The advantage of this rather remarkable result is that the problem may be solved for significantly longer interconnects. Here we restrict our attention to $L=50 \mu\text{m}$ and so are able to continue with $\varepsilon=0.1 \mu\text{m}$ and around 500 mesh points per interconnect, however taking $\varepsilon=0.5 \mu\text{m}$ and around 1000 mesh points one could consider $500 \mu\text{m}$ lines.

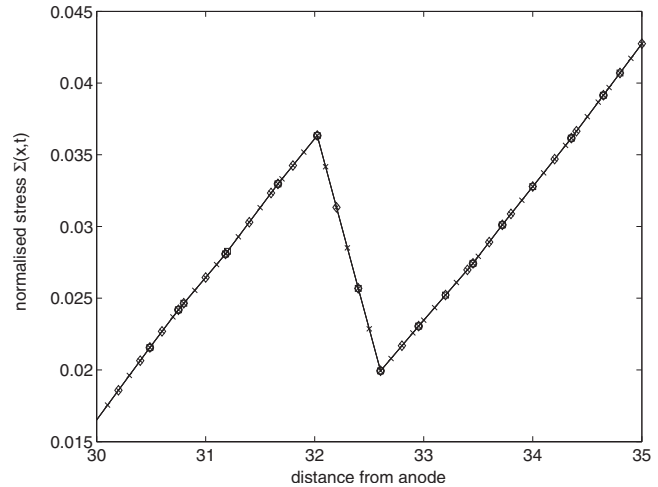


FIG. 4. Close up of one region of Fig. 3.

Equation (24) has been solved for the void nucleation time on 10^5 lines constructed with grain diameters distributed lognormally [$d_{50}=0.5 \mu\text{m}$ and $\sigma_d=0.36$ (Ref. 38)], and an effective diffusivity with an activation energy drawn from a normal mixture (20% slow grains, 80% fast). This gives a lognormal mixture for κ_{eff} of two distributions with a fast/slow ratio of 8 and lognormal standard deviation of 0.9 for both diffusion speeds. Z^* and hence G and α are taken to be independent of position. The choices for the parameters defining the distribution of effective diffusivities are fairly random as insufficient data on such parameter is available. The split between slow and fast diffusing grains is in line with Ref. 34 the other parameters above must be regarded as unknown, so that several simulations have been run for a variety of values. However, for the case described, the CDF looks reassuringly relatively straight on a lognormal plot, Fig. 5. The remaining parameters for this simulation are $L=50 \mu\text{m}$, $\Omega=1.16 \times 10^{-29} \text{m}^3$, $T=300 \text{ }^\circ\text{C}$, $\sigma_{\text{cr}}=50 \text{ MPa}$, $Z^*=0.7$, $j=20 \text{ mA}/\mu\text{m}$, $\rho=2.1 \times 10^{-8} \Omega\text{m}$, $B \sim 28 \text{ GPa}$, and $\alpha=Z^*q\rho j/kT \sim 0.0059 \mu\text{m}^{-1}$.⁴²⁻⁴⁵

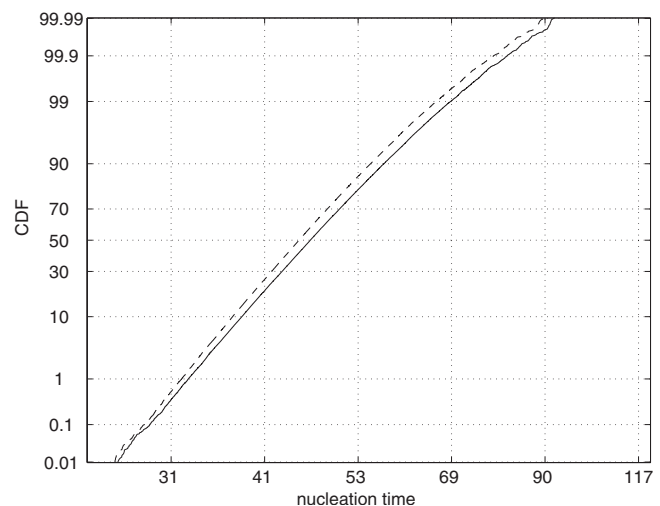


FIG. 5. The nucleation time distribution (solid curve) for linear model (not lumped) compared to that obtained by a two-exponential lumped, MoM approximation (dashed-curve). $L=50 \mu\text{m}$ and copper parameters as indicated in text. 10^5 separate interconnects are simulated.

A void will nucleate at the cathode via when the stress reaches the critical value σ_{cr} , consequently the nucleation time is obtained from the solution to $c_{N+1}(t_{nuc}) = \Omega \sigma_{cr} / k_B T$. Constructing the matrix exponential in Eq. (24) involves the diagonalization of the matrix $\mathbf{M}^{-1}\mathbf{K}$ and serves to introduce another approximation, known as mass matrix lumping, which is standard in finite element methods. Although not altogether necessary here, such lumping will be useful in attempting some analytical solutions later, and extending the analysis to longer interconnects. Lumping the mass matrix (\mathbf{M} here) means approximating \mathbf{M} by $\hat{\mathbf{M}}$, a diagonal matrix, whose diagonal element in row k is the k^{th} row sum of \mathbf{M} . In the present case this means: $\hat{\mathbf{M}}(1,1) = \varepsilon_1/2$, $\hat{\mathbf{M}}(2,2) = \varepsilon_1/2 + \varepsilon_2/2$, $\hat{\mathbf{M}}(k,k) = \varepsilon_{k-1}/2 + \varepsilon_k/2$ ($k=2, \dots, N$), and $\hat{\mathbf{M}}(N+1, N+1) = \varepsilon_N/2$. The utility of this approximation is that it allows $\hat{\mathbf{M}}$ to be inverted trivially in Eq. (24). However even with such mass lumping; the resulting tridiagonal nature of $\mathbf{M}^{-1}\mathbf{K}$ and the simplicity in diagonalizing such matrices,⁴⁶ evaluating $\exp(\mathbf{M}^{-1}\mathbf{K}t)$ at sufficient times to estimate a failure time is a difficult task, involving the diagonalization (diagonal matrix \mathbf{D} and matrix of eigenvectors \mathbf{S}) the product $\mathbf{S}^{-1} \exp(\mathbf{D}t)\mathbf{S}$ at each time t . Progress can be made as interestingly the mass lumping makes the stress development on the mesh isomorphic to the development of a pseudovoltage on a lumped-element CR transmission line, in which the capacitance per unit length is constant and the local line resistivity is proportional to the grain diffusivity. The interest in this is that calculating time constants for CR trees or networks is already commonplace in static timing analysis (STA) (Refs. 47–49) for very large scale integrated (VLSI) circuits and, as system reliability on interconnect nets is of fundamental interest in EM, VLSI timing techniques may be helpful. The equivalent CR model, with a lumped mass matrix, may be obtained directly from Eq. (23) or in the following manner.

To simplify the analysis we once again keep the G_k , and hence α_k , values constant (largely equivalent to keeping the Z_k^* values constant). Linearizing the system, and making the substitution $V(x,t) = \Sigma(x,t) - \alpha x = \Sigma(x,t) - \Sigma(x,t \rightarrow \infty)$, Eq. (3) can be integrated over the region between two neighboring mesh points (x_j, x_{j+1}) (within which $\kappa = \kappa_{j+1}$ is constant) to obtain

$$\begin{aligned} V(x_{j+1}, t) - V(x_j, t) &= -\frac{1}{\kappa_{j+1}} \int_{x_j}^{x_{j+1}} J(x, t) dx \\ &\approx -\frac{(x_{j+1} - x_j)}{\kappa_{j+1}} J\left(\frac{x_{j+1} + x_j}{2}, t\right). \end{aligned} \quad (25)$$

If the mesh is taken sufficiently small then J varies roughly linearly with x over the mesh element then the approximation in Eq. (25) is exact. Likewise integration of Eq. (5) gives

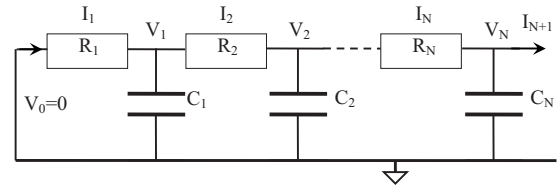


FIG. 6. A lumped element model of the interconnect. Each mesh interval is approximated by a C–R–C π -element. For most calculations here an element is of length of around $d_{50}/5$.

$$\begin{aligned} J\left(\frac{x_{j+1} + x_j}{2}, t\right) - J\left(\frac{x_j + x_{j-1}}{2}, t\right) &= \frac{\partial}{\partial t} \int_{(x_j + x_{j-1})/2}^{(x_{j+1} + x_j)/2} V(x, t) dx \\ &\approx -\frac{(x_{j+1} - x_{j-1})}{2} \frac{\partial}{\partial t} V(x_j, t), \end{aligned} \quad (26)$$

where the approximation is again valid if the mesh is sufficiently small. In the limit of an infinitesimal mesh there is of course no approximation. This allows for the accuracy of these approximations to be estimated by solving the problem on several different sized meshes, if the solution has converged, as is the case here (Figs. 3 and 4), then one can be reasonably confident in the approximation.

If we treat J as a pseudocurrent and V as a pseudovoltage then it is clear that Eq. (25) describes the behavior of a resistor in an equivalent circuit and Eq. (26) describes that of a capacitor, Fig. 6. To illustrate the point we consider an interconnect with parameters appropriate to copper, with a pad at $x=0$ and a blocking via at $x=L$. In this case the end conditions are $\Sigma(x=0, t) = V(x=0, t) = 0$ and $J(x=L, t) = 0$. If the initial stress $\sigma(x, t=0) = 0$ then the initial pseudovoltage is $V(x, t=0) = -\alpha x$. Each mesh interval $\varepsilon_{j+1} = x_{j+1} - x_j$ is effectively represented by a C–R–C, π -element according to Eqs. (25) and (26), Fig. 6. If we eliminate the currents between Eqs. (25) and (26) we obtain

$$\begin{aligned} \frac{\varepsilon_{j+1} + \varepsilon_j}{2} \frac{d}{dt} V(x_j, t) &= -\frac{V(x_{j+1}, t) - V(x_j, t)}{\varepsilon_j / \kappa_j} \\ &\quad + \frac{V(x_j, t) - V(x_{j-1}, t)}{\varepsilon_{j-1} / \kappa_{j-1}} \end{aligned} \quad (27)$$

or in matrix notation

$$\begin{aligned} \hat{\mathbf{M}} \frac{d}{dt} \mathbf{V}(t) &= -\mathbf{K} \mathbf{V}(t), \\ \hat{\mathbf{M}} \frac{d}{dt} \mathbf{c}(t) &= -\mathbf{K} \mathbf{c}(t) + \mathbf{K} \mathbf{c}_\infty, \end{aligned} \quad (28)$$

i.e., Eq. (23) again, after mass lumping. Thus, within the 1D model of Korhonen *et al.*, the stress development at the grain boundary end points on the copper/nitride-cap interface mimics that of a pseudovoltage on a simple passive CR network. The reason for stressing this isomorphism is that there is a very large body of work on such circuits as an approximation of the transistor networks in STA for VLSI designs. Whereas here we are interested in the time that the stress at x_k reaches the critical value of σ_{cr} for nucleation, there the interest is in the time that the pseudovoltage takes to reach the transistor threshold voltage, V_{Th} . If we pursue things a

little further, Eq. (28) may be solved using normal methods, giving a solution represented by a sum of exponential terms whose exponents are determined by the eigenvalues of the matrix $\Xi^{-1} = \mathbf{M}^{-1}\mathbf{K}$. Likewise $c_k(t)$ may be expressed as a sum of $N+1$ exponentials, i.e.,

$$c_k(t) = (\mathbf{c}_\infty)_k + \sum_{k=1}^{N+1} A_{k,n} \exp(-\lambda_n t). \quad (29)$$

Clearly, at large times, which also corresponds to the case of short lines, the series in Eq. (29) may be truncated after the dominant exponential term. That is

$$\Sigma(x_k, t) = c_k(t) \approx (\mathbf{c}_\infty)_k + A_k \exp(-\lambda_{\min} t), \quad (30)$$

and the nucleation time t_{nuc1} is given approximately by

$$t_{\text{nuc1}} \approx -\frac{1}{\lambda_{\min}} \log\left(\frac{\Sigma_{\text{cr}} - \alpha L}{A_L}\right). \quad (31)$$

The time-constant $1/\lambda_{\min}$ is known in STA as the (first) Elmore delay.⁴⁷ Here we note that if A_L and λ_{\min} can be calculated analytically, it may be possible to obtain some (analytical) idea of the void nucleation time distribution. Consequently we continue a little with this analogy.

IV. AN ANALYTICAL CONSIDERATION OF THE VOID NUCLEATION TIME DISTRIBUTION

The linearized version of Eqs. (3) and (5) may be solved exactly for the homogeneous line using standard Fourier-Laplace methods,^{33,50} when the stress is naturally represented by an infinite sum of exponential terms. With the initial and boundary conditions considered above one easily obtains

$$\begin{aligned} \Sigma(x, t) = \alpha x - \frac{8\alpha L}{\pi^2} \sum_{n=0}^{\infty} \frac{\sin\left[\frac{(2n+1)\pi x}{2L}\right]}{(2n+1)^2} \\ \times \exp\left[-\pi^2 \left(\frac{2n+1}{2}\right)^2 \frac{\kappa t}{L^2}\right]. \end{aligned} \quad (32)$$

In short lines ($L < 50 \mu\text{m}$ in copper) the first exponential term dominates at all but very small times. At the cathode $x=L$ we may thus approximate

$$\Sigma(x=L, t) - \alpha L \approx -\frac{8\alpha L}{\pi^2} \exp\left(-\frac{\pi^2 \kappa t}{4L^2}\right), \quad (33)$$

corresponding to an exponential pseudovoltage of $V(x=L, t) \sim -8\alpha L \exp(-\pi^2 \kappa t / 4L^2) / \pi^2$. This result can now be used as a useful check of the result for inhomogeneous lines in that from the solutions in Eqs. (29) and (30) we should regain Eqs. (32) and (33), in the limit of $\varepsilon \rightarrow 0$ and with all $\kappa_k = \kappa$ (constant).

In the case of an inhomogeneous line we can proceed in the following manner, summing Eq. (26) over all mesh points from k to N gives

$$0 - J\left(\frac{x_k + x_{k-1}}{2}, t\right) = -\sum_{j=k}^N \frac{(x_{j+1} + x_{j-1})}{2} \frac{\partial}{\partial t} V(x_j, t), \quad (34)$$

and, summing Eq. (25) over all k gives, and substituting from Eq. (34) leads to

$$\begin{aligned} V(x = x_{N+1} = L, t) \\ = -\sum_{k=0}^N \frac{(x_{k+1} - x_k)}{\kappa_k} \sum_{j=k}^N \frac{(x_{j+1} + x_{j-1})}{2} \frac{\partial}{\partial t} V(x_j, t). \end{aligned} \quad (35)$$

If we assume that for short lines ($L < 50 \mu\text{m}$) the inhomogeneous line solution will also be dominated by a single exponential from the series solution in Eq. (35), we expect that $V(x_k, t) \sim -A_k \exp(-t/\tau_k)$ for some A_k and for some τ_k . Using Eq. (35) and some method of moments (MoM) it is possible to estimate the values of A_{N+1} and τ_{N+1} and then compare the results to the equivalent values appearing in Eq. (33). To do this integrate Eq. (35) over all t to get the first two moments of $V(L, t)$ in terms of known parameters as

$$\begin{aligned} \mathbf{m}_0^{(N+1)} &\equiv \int_0^\infty V(L, t) dt \\ &= -\alpha \sum_{k=0}^N \frac{(x_{k+1} - x_k)}{\kappa_k} \sum_{j=k}^N \frac{(x_{j+1} + x_{j-1})}{2} x_j, \\ \mathbf{m}_1^{(N+1)} &\equiv \int_0^\infty t V(L, t) dt \\ &= -\alpha \sum_{k=0}^N \frac{(x_{k+1} - x_k)}{\kappa_k} \sum_{j=k}^N \frac{(x_{j+1} + x_{j-1})}{2} \\ &\quad \times \int_0^\infty V(x_k, t) dt, \end{aligned} \quad (36)$$

Now A_{N+1} and τ_{N+1} may be estimated by inserting $V(x_{N+1}, t) \sim -A_{N+1} \exp(-t/\tau_{N+1})$ into Eq. (36). Solving gives $\tau_{N+1} = \mathbf{m}_1^{(N+1)} / \mathbf{m}_0^{(N+1)}$ and $A_{N+1} = -\mathbf{m}_0^{(N+1)} / \tau_{N+1} = -(\mathbf{m}_0^{(N+1)})^2 / \mathbf{m}_1^{(N+1)}$. Taking κ to be constant in Eq. (36) should lead to the solution for the homogeneous line, and proceeding to the limit (mesh size $\varepsilon \rightarrow 0$) we obtain $A_{N+1} = 5\alpha L/6$ and $\tau_{N+1} = L^2/2.5\kappa$. These are very close (-2.81% and $+1.30\%$) to the exact results of $A = 8\alpha L/\pi^2$ and $\tau = 4L^2/\pi^2\kappa$ in Eq. (33). The error between the two comes from the mass lumping and indicates a rough size of that effect. Defining the vector \mathbf{X}_k of mesh positions (x_k) note that, in general, the vector of first moments at the mesh points x_k , is $\mathbf{m}_0 = \alpha \mathbf{K}^{-1} \mathbf{M} \mathbf{X} = \alpha \Xi \mathbf{X}$ and $\mathbf{m}_1 = \mathbf{K}^{-1} \mathbf{M} \mathbf{m}_0 = \alpha \Xi^2 \mathbf{X}$. Thus within a one exponential approximation we have, for the nucleation time,

$$\begin{aligned} t_{\text{nuc1}} = \tau_{N+1} \log\left(\frac{A_{N+1}}{\alpha L - \Sigma_{\text{cr}}}\right) = \frac{\mathbf{m}_1^{(N+1)}}{\mathbf{m}_0^{(N+1)}} [2 \log(\mathbf{m}_0^{(N+1)}) \\ - \log(\mathbf{m}_1^{(N+1)}) - \log(\alpha L - \Sigma_{\text{cr}})], \end{aligned} \quad (37)$$

and an analysis of the statistical properties of \mathbf{m}_0 and \mathbf{m}_1 should lead to those of t_{nuc1} . The value of the time constant of the pseudovoltage at $x=L$, τ_{N+1} , is known as the (first) Elmore delay in STA.⁴⁷

In point of fact for $50 \mu\text{m}$ lines, the nucleation times predicted by Eq. (37) agree with those predicted by Eq. (24) only for relatively long failure times (above the t_{50} for the system). More accurate estimates of t_{nuc1} can be obtained by using more exponentials in the approximation of $V(L, t)$. This

corresponds to a higher order MoM approximation^{48,49} which is the standard procedure in STA. In addition rigorous bounds on the delay in an RC network may also be obtained which should then place rigorous bounds⁴⁸ on the nucleation time predicted by the SEM model. To do thus define the moments \mathbf{m}_n of the pseudovoltage vector $[\mathbf{V}_k(t) = \mathbf{V}(x_k, t)]$ in the following manner. Let us denote $\mathbf{K}^{-1}\mathbf{M} = \mathbf{\Xi}$ when the linear stress evolution equation Eq. (28) may be written as

$$\mathbf{V}(t) = -\mathbf{K}^{-1}\mathbf{M}\frac{d}{dt}\mathbf{V}(t) = -\mathbf{\Xi}\frac{d}{dt}\mathbf{V}(t). \quad (38)$$

Then the vector \mathbf{m}_n of n^{th} moments of $\mathbf{V}(x_k, t)$ is defined as

$$\mathbf{m}_n \equiv \int_0^\infty \frac{t^n}{n!} \mathbf{V}(t) dt = \mathbf{\Xi} \mathbf{m}_{n-1}, \quad (39)$$

and, by recursion, $\mathbf{m}_n = \mathbf{\Xi}^n \mathbf{m}_0$. Higher order matching techniques, which include more exponentials in the series solution equivalent to Eq. (32), may now be used. A q th-order model, using the $2q$ moments \mathbf{m}_0 to \mathbf{m}_{2q-1} , will match the first q exponentials in the series (matching to an exponent and a pre-exponential multiplier).⁴⁹ However to make analysis tractable we wish to keep this order as low as possible and it appears that two exponentials are generally enough.

A two exponential model ($q=2$) represents the pseudovoltage at $x=x_k$ as

$$\mathbf{V}_k(t) = \mathbf{A}_k \exp(-t/\tau_{1k}) + \mathbf{B}_k \exp(-t/\tau_{2k}), \quad (40)$$

The first four moments of this ($n=0, 1, 2, 3$) are then given by

$$\mathbf{m}_n^{(k)} = \int_0^\infty \frac{t^n}{n!} \mathbf{V}_k(t) dt = \mathbf{A}_k \tau_{1k}^{n+1} + \mathbf{B}_k \tau_{2k}^{n+1}. \quad (41)$$

Eliminating \mathbf{A}_k and \mathbf{B}_k from Eq. (41), and rearranging, the time exponents τ_{1k} and τ_{2k} are found to be given by the roots of the quadratic

$$\begin{aligned} (\mathbf{m}_1^{(k)2} - \mathbf{m}_2^{(k)} \mathbf{m}_0^{(k)}) \tau_k^2 + (\mathbf{m}_3^{(k)} \mathbf{m}_0^{(k)} - \mathbf{m}_2^{(k)} \mathbf{m}_1^{(k)}) \tau_k + (\mathbf{m}_2^{(k)2} \\ - \mathbf{m}_3^{(k)} \mathbf{m}_1^{(k)}) = 0, \end{aligned} \quad (42)$$

and the pre-exponential multipliers are obtained from Eq. (41) as

$$\begin{pmatrix} \mathbf{A}_k \\ \mathbf{B}_k \end{pmatrix} = \begin{bmatrix} \tau_{1k} & \tau_{2k} \\ \tau_{1k}^2 & \tau_{2k}^2 \end{bmatrix}^{-1} \begin{pmatrix} \mathbf{m}_0^{(k)} \\ \mathbf{m}_1^{(k)} \end{pmatrix}. \quad (43)$$

Finally obtaining these moments first from $\mathbf{m}_0 = \alpha \mathbf{K}^{-1} \mathbf{M} \mathbf{X} = \alpha \mathbf{\Xi} \mathbf{X}$, and then recursively from $\mathbf{m}_n = \mathbf{\Xi}^n \mathbf{m}_0$, the pseudovoltage and hence the normalized stress $\Sigma(L, t)$ is easily obtained. Figure 5 shows a comparison of the CDF for 10^5 interconnect via nucleation times obtained by solving Eq. (28) (solid line) and from solving the two-exponential approximation, Eqs. (40)–(43) (dotted line). The grain size distribution is taken as lognormal and diffusion activation energy distribution is formed from the normal mixture of fast and slow values. It is clear that the distribution of failure-times from Eqs. (40) and (28) are reasonably straight when plotted on lognormal probability paper and that the two exponential model is a good approximation to the “true” nucleation time, Fig. 7. Such an approximation replaces the need

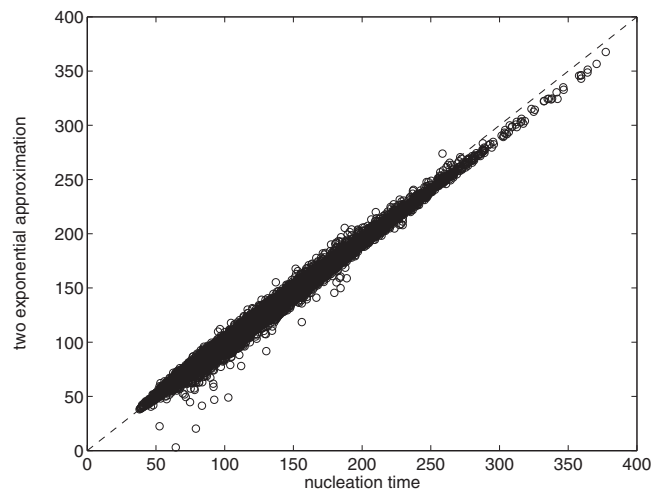


FIG. 7. Same as Fig. 5. The “true” nucleation time, solving the linear (not lumped) equation, Eq. (24), plotted against the two-exponential approximation Eq. (40).

to both diagonalize the matrix $\mathbf{M}^{-1}\mathbf{K}$ in Eq. (24) and to find $\mathbf{S}^{-1} \exp(\mathbf{D}t)\mathbf{S}$ at a large number of time values, and potentially high-order MoM approximations may be used for longer lines.

V. A LOGNORMAL DISTRIBUTION

The question that we set out to answer in this work is, what sort of void nucleation time distribution is predicted by assuming the SEM of Korhonen *et al.*? To answer this we consider the one exponential approximation of Eq. (30). Even though this is not a good approximation for lines of length $50 \mu\text{m}$ it is perfectly adequate for lines of $30 \mu\text{m}$ and Eq. (30) presents far easier analysis than Eq. (40). For line lengths around $30 \mu\text{m}$, the nucleation time t_{nucl} is given by Eq. (37) to a reasonable accuracy so its statistical properties depend upon those of $\mathbf{m}_0^{(N+1)}$ and $\mathbf{m}_1^{(N+1)}$, which in turn given by Eqs. (36). Rewriting Eq. (36) with all grains of equal size d , say, we obtain

$$\begin{aligned} \mathbf{m}_0^{(N+1)} &= -\frac{\alpha d^2}{6} \sum_{k=0}^N \frac{N(2N+1)(N+1) - k(2k+1)(k+1)}{\kappa_k} \\ &= -\frac{\alpha d^2}{3} \sum_{k=0}^N \frac{N^3 - k^3}{\kappa_k}. \end{aligned} \quad (44)$$

If the effective diffusivities κ_k are drawn from a lognormal mixture, this corresponds to a (weighted) sum of lognormal terms. The sum will be dominated by the relatively small number of slow grains as these have the largest $1/\kappa_k$ values, meaning that $\mathbf{m}_0^{(N+1)}$ is the sum of a relatively small number of lognormal terms. However the permanence property^{51–54} of the lognormal means that sums (and weighted sums) of lognormals are also roughly lognormal. Because of their skew the CLT does not apply until there are a very large number of lognormal terms. Allowing the grain sizes to vary does not alter this argument significantly. In a similar manner it may be shown that $\mathbf{m}_1^{(N+1)}$ will also be roughly lognormal and hence also $\mathbf{m}_1^{(N+1)}/\mathbf{m}_0^{(N+1)}$. This would also mean that $\log(\mathbf{m}_0^{(N+1)})$ and $\log(\mathbf{m}_1^{(N+1)})$ are normally distributed. The

nucleation time given in Eq. (37) is, consequently, roughly the product of a lognormal variable $\mathbf{m}_1^{(N+1)}/\mathbf{m}_0^{(N+1)}$ and a normal variable $2 \log(\mathbf{m}_0^{(N+1)}) - \log(\mathbf{m}_1^{(N+1)}) - \log(\alpha L - \Sigma_{cr})$. Such a product of independent variables, to an extent that depends on the variance of the elements, can have shape which is between the symmetric normal and the skewed lognormal. However because the normal and lognormal variates come from the same source, i.e., $\mathbf{m}_1^{(N+1)}$ and $\mathbf{m}_0^{(N+1)}$, the elements are not independent and one would expect the lognormal's skewness to dominate. We can conclude that the nucleation time is roughly a weighted sum of lognormal terms provided that the activation energy for diffusion is normally distributed; that that sum is dominated by a small number of slow grains and that the permanence property of lognormals prevents the CLT from destroying the lognormality. As a result this approximation to t_{nuc} should be distributed roughly lognormally.

In point of fact, although the single-exponential distribution shows up quite well as a lognormal even for 50 μm lines, and fits the high end of the distribution very well, the important early-failure tail of the distribution is poorly represented; the two-exponential solution of Eq. (40) provides a much more faithful fit, Figs. 5 and 7. Figure 7 shows the two solutions on a lognormal plot showing both distributions to be close to lognormal. The sort of theoretical analysis presented above is not straightforward for the two-exponential approximation, but one can probably expect the same sort of behavior, i.e., if the Diffusivities are lognormal the nucleation time will be close also.

More important than a possible explanation of lognormality though, is a framework for investigating the distribution of t_{nuc} in other cases, and between processes.

VI. SUMMARY AND CONCLUSIONS

In this paper we have attempted to model the distribution for void nucleation times under EM in copper IC interconnects. Assuming that the stress distribution evolves according to the 1D model of Korhonen *et al.*, and that the lines have a bamboo microstructure with transverse grains, we are able to calculate the build-up of hydrostatic stress in a straightforward manner. Using finite element methods, we have demonstrated two important results. First, it is valid to ignore the dependence of diffusivity on the tensile stress, and second, that a finite element mesh which only contains the grain boundary points ($\mathcal{N}=\mathcal{G}$) gives accurate results.

Modeling the lines as a 1D string of bamboo grains with a lognormal size distribution (d) and assigning effective diffusivities with a normal mixture of effective activation energies (E_A), the nucleation time for voiding at the cathode via may be rapidly calculated exactly using standard MATLAB matrix methods. For a variety of sensible parameter values the resulting distribution looks reassuringly straight on a lognormal plot.

In addition we have attempted to obtain information about the nucleation time distribution by analytical methods. For short lines (up to around 50 μm), mass-lumping replaces the finite element model by an equivalent RC circuit, from which the dominant time constant may be easily ex-

tracted. Its dominance occurs as a short line length maps on to long time behavior and vice versa. MoM type fitting (q exponentials in the series requires the moments \mathbf{m}_0 to \mathbf{m}_{2q-1}) can provide a good approximation to the nucleation time which allows some further analytical investigation of its distribution. A simple analysis shows that there may be good reasons for the use of a lognormal distribution when the diffusion activation energy is distributed according to a normal mixture.

The simplicity of this analysis will allow a number of other problems to be considered will relative ease. For example, the nucleation time in simple tree networks, in a single layer of metallization, may be analyzed in a manner which allows the grain diffusivities in different limbs to be drawn from the same or from different distributions. Alternatively one may investigate the sorts of distribution of D , d , or σ_{cr} values that would give rise to the observed occurrence of in line void nucleation. Patching together two solutions up-stream and down-stream some void dynamics may also be modeled.

- ¹T. Berger, L. Arnaud, R. Gonella, G. Lormand, and Y. Morand, *Microelectron. Reliab.* **40**, 1311 (2000).
- ²Reliability in CMOS IC design: Physical failure mechanisms and their modelling, *in* MOSIS Technical Notes, <http://www.mosis.org/support/technicalnotes.html>.
- ³B. Li, T. D. Sullivan, T. C. Lee, and D. Badami, *Microelectron. Reliab.* **44**, 365 (2004).
- ⁴C. J. Christensen, B. Li, J. Gill, R. Filippi, and M. Angyal, *IEEE Trans. Device Mater. Reliab.* **6**, 163 (2006).
- ⁵M. J. Attardo, R. Rutledge, and R. C. Jack, *J. Appl. Phys.* **42**, 4343 (1971).
- ⁶D. J. LaCombe and E. L. Parks, *IEEE Int. Reliab. Phys. Symp. Proc.* **24**, 1 (1986).
- ⁷J. R. Lloyd, *Microelectron. Reliab.* **47**, 1468 (2007).
- ⁸V. Sukharev, E. Zschech, and W. D. Nix, *J. Appl. Phys.* **102**, 053505 (2007).
- ⁹J. R. Black, *Proc. IEEE* **57**, 1587 (1969).
- ¹⁰J. R. Black, *IEEE Int. Reliab. Phys. Symp. Proc.* **6**, 148 (1966).
- ¹¹D. B. Kececioglu, *Reliability Engineering Handbook* (Prentice Hall, Englewood Cliffs, 1991), Vol. 1.
- ¹²J. R. Lloyd, *J. Appl. Phys.* **69**, 7601 (1991).
- ¹³H. A. Schafft, T. C. Grant, A. N. Saxena, and C.-Y. Kao, *IEEE Int. Reliab. Phys. Symp. Proc.* **23**, 93 (1985).
- ¹⁴P. B. Gbate, *IEEE Int. Reliab. Phys. Symp. Proc.* **20**, 292 (1982).
- ¹⁵J. C. Blair, P. B. Gbate, and C. T. Haywood, *Appl. Phys. Lett.* **17**, 281 (1970).
- ¹⁶V. B. Fiks, *Sov. Phys. Solid State* **1**, 14 (1959).
- ¹⁷H. B. Huntington and A. R. Grone, *J. Phys. Chem. Solids* **20**, 76 (1961).
- ¹⁸M. Shatzkes and J. R. Lloyd, *J. Appl. Phys.* **59**, 3890 (1986).
- ¹⁹R. Kirchheim, *Acta Metall. Mater.* **40**, 309 (1992).
- ²⁰M. A. Korhonen, P. Borgesen, K. N. Tu, and C.-Y. Li, *J. Appl. Phys.* **73**, 3790 (1993).
- ²¹V. M. Dwyer, F.-S. Wang, and P. Donaldson, *J. Appl. Phys.* **76**, 7305 (1994).
- ²²J. J. Clement and C. V. Thompson, *J. Appl. Phys.* **78**, 900 (1995).
- ²³S. M. Alam, C. L. Gan, C. V. Thompson, and D. E. Troxel, *Microelectron. J.* **38**, 463 (2007).
- ²⁴S. P. Hau-Riege, *J. Appl. Phys.* **91**, 2014 (2002).
- ²⁵I. A. Blech, *J. Appl. Phys.* **47**, 1203 (1976).
- ²⁶I. A. Blech and C. Herring, *Appl. Phys. Lett.* **29**, 131 (1976).
- ²⁷Z.-S. Choi, R. Mönig, and C. V. Thompson, *Appl. Phys. Lett.* **90**, 241913 (2007).
- ²⁸Z.-S. Choi, R. Mönig, and C. V. Thompson, *J. Mater. Res.* **23**, 383 (2008).
- ²⁹M. Hauschildt, M. Gall, S. Thrasher, P. Justison, L. Michaelson, R. Hernandez, H. Kawasaki, and P. S. Ho, in *Seventh International Workshop on Stress-Induced Phenomena in Metallization*, AIP Conference Proceedings Vol. 741, edited by P. S. Ho, S. P. Baker, T. Nakamura, and C. A. Volkert (American Institute of Physics, Melville, NY, 2004), p. 112.
- ³⁰M. Hauschildt, M. Gall, and R. Hernandez, in *Tenth International Work-*

- shop on Stress-Induced Phenomena in Metallization*, AIP Conference Proceedings Vol. 1143, edited by P. S. Ho, E. Zschech, and E. Ogawa (American Institute of Physics, Melville, NY, 2009), p. 31.
- ³¹M. Hauschildt, Ph.D. thesis, 2005; Available online at www.lib.utexas.edu/etd/d/2005/hauschildtm39810/hauschildtm39810.pdf.
- ³²L. M. Gignac, C.-K. Hu, and E. G. Liniger, *Microelectron. Eng.* **70**, 398 (2003).
- ³³V. M. Dwyer, *J. Appl. Phys.* **104**, 053708 (2008).
- ³⁴C. Ryu, A. L. S. Loke, T. Nogami, and S. S. Wong, *IEEE Int. Reliab. Phys. Symp. Proc.* **35**, 201 (1997).
- ³⁵R. J. Gleixner and W. D. Nix, *J. Appl. Phys.* **83**, 3595 (1998).
- ³⁶V. M. Dwyer, *J. Appl. Phys.* **96**, 2914 (2004).
- ³⁷B. D. Knowlton and C. V. Thompson, *J. Mater. Res.* **13**, 1164 (1998).
- ³⁸M. W. Lane, E. G. Liniger, and J. R. Lloyd, *J. Appl. Phys.* **93**, 1417 (2003).
- ³⁹J. R. Lloyd, M. W. Lane, E. G. Liniger, C.-K. Hu, T. M. Shaw, and R. Rosenberg, *IEEE Trans. Device Mater. Reliab.* **5**, 113 (2005).
- ⁴⁰L. Lapidus and G. F. Pinder, *Numerical Solutions of Partial Differential Equations in Science and Engineering* (Wiley, New York, 1982).
- ⁴¹L. F. Shampine and M. W. Reichelt, *SIAM J. Sci. Comput. (USA)* **18**, 1 (1997).
- ⁴²L. Arnaud, G. Taravel, T. Berger, D. Mariolle, Y. Gobil, and I. Touet, *Microelectron. Reliab.* **40**, 77 (2000).
- ⁴³C. S. Hau Riege, S. P. Hau Riege, and A. Marathe, *J. Appl. Phys.* **96**, 5792 (2004).
- ⁴⁴E. T. Ogawa, K.-D. Lee, V. A. Blaschke, and P. S. Ho, *IEEE Trans. Reliab.* **51**, 403 (2002).
- ⁴⁵J. He and Z. Suo, Proceedings of the International Workshop on Stress-Induced Phenom in Metal, 2004 (unpublished), Vol. 7, p. 1.
- ⁴⁶J. K. Cullam and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*, Classics in Applied Mathematics Vol. 1 (SIAM, Philadelphia, PA, 2002).
- ⁴⁷W. C. Elmore, *J. Appl. Phys.* **19**, 55 (1948).
- ⁴⁸J. Rubinstein, P. Penfield, Jr., and M. A. Horowitz, *IEEE Trans. C-AD* **2**, 202 (1983).
- ⁴⁹L. T. Pillage and R. A. Rohrer, *IEEE Trans. C-AD* **9**, 352 (1990).
- ⁵⁰J. Crank, *The Mathematics of Diffusion*, 2nd ed. (Oxford, UK, 1975).
- ⁵¹R. L. Fante, *IEEE Trans. Aerosp. Electron. Syst.* **37**, 739 (2001).
- ⁵²R. Barakat, *J. Opt. Soc. Am.* **66**, 211 (1976).
- ⁵³N. B. Mehta, J. Wu, A. F. Molisch, and J. Zhang, *IEEE Trans. Wirel. Comm.* **6**, 2690 (2007).
- ⁵⁴N. C. Beaulieu and F. Rajwani, *IEEE Commun. Lett.* **8**, 709 (2004).