The influence of microstructure on the probability of early failure in aluminum-based interconnects

V. M. Dwyer[a)]
Department of Electronic and Electrical Engineering, Loughborough University, Loughborough LE11 3TU, United Kingdom

(Received 1 March 2004; accepted 20 May 2004)

For electromigration in short aluminum interconnects terminated by tungsten vias, the well known "short-line" effect applies. In a similar manner, for longer lines, early failure is determined by a critical value L_{crit} for the length of polygranular clusters. Any cluster shorter than L_{crit} is "immortal" on the time scale of early failure where the figure of merit is not the standard t_{50} value (the time to 50% failures), but rather the total probability of early failure, P_{cf}. P_{cf} is a complex function of current density, linewidth, line length, and material properties (the median grain size d_{50} and grain size shape factor σ_d). It is calculated here using a model based around the theory of runs, which has proved itself to be a useful tool for assessing the probability of extreme events. Our analysis shows that P_{cf} is strongly dependent on σ_d, and a change in σ_d from 0.27 to 0.5 can cause an order of magnitude increase in P_{cf} under typical test conditions. This has implications for the web-based two-dimensional grain-growth simulator MIT/EmSim, which generates grain patterns with $\sigma_d = 0.27$, while typical as-patterned structures are better represented by a σ_d in the range 0.4 – 0.6. The simulator will consequently overestimate interconnect reliability due to this particular electromigration failure mode. © *2004 American Institute of Physics*. [DOI: 10.1063/1.1771825]

I. INTRODUCTION

Electromigration-induced early failure of metallization represents a significant problem for the accurate analysis of modern integrated circuit (IC) reliability. However, these early failures (i) may occur with such small probability that they can be missed in standard sized electromigration tests (or at least regarded as outliers); (ii) may not obey the same statistics as those longer time failures which are observed in tests; and (iii) may not indeed even be caused by the same failure mode. Clearly, a proper appreciation of the extent of early failures in accelerated tests is vital in order to ensure the correct extrapolation to operational conditions. In particular, it is important to know the dependence of the early failure probability on acceleration parameters.

The short line effect is a well known feature of aluminum-based interconnect. For a given applied current density j, critical line lengths exist for both void nucleation $L_B^{(n)}$ (Refs. 1–3) and void growth failure $L_B^{(g)}$.[4,5] Below these critical lengths a stress-induced back-flux cuts off the metal migration either before the void has nucleated or before it has achieved sufficient growth to cause the circuit to malfunction. Above both, failure will occur at a time dependent upon the diffusivity along the line (e.g., Ref. 5). In fine line metallization most of the line will be made from spanning grains,[6–8] giving its microstructure an appearance similar to bamboo, Fig 1. However, there will also be regions of the line containing grain boundaries running along the line length.[6–8] Where several of these grain boundaries join together they define polygranular segments or clusters, Fig. 1(b). Despite the fact that grain boundary diffusion only takes place within the narrow boundary at the intersection of two grains $\delta(\sim 1$ nm) while transgranular diffusion occurs across the entire linewidth w, the effective diffusion within the clusters is significantly greater than that through spanning (or bamboo) grains.[6–8] As a result the bamboo grains at the ends of a cluster act as blocking boundaries and perform the same role that vias do for short line failure. A version of the short line effect consequently operates for clusters.[9] The greater diffusivity in a cluster means that, in general, a void nucleating at a cluster end will occur more rapidly than a void nucleating at the cathode via. The result is that the presence of polygranular segments, or clusters, represents a possible cause of early failures in near bamboo interconnects. The purpose of this paper is to analyze the contribu-

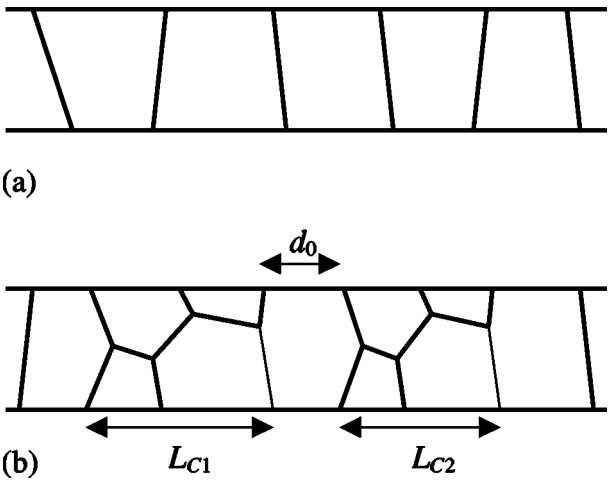

FIG. 1. Polygranular clusters in a near-bamboo line. (a) shows a schematic of a pure bamboo region consisting entirely of spanning or bamboo grains, while (b) shows a schematic of a region containing two clusters of lengths L_{C1} and L_{C2} separated by a single bamboo grain of grain size d_0.

a)Electronic mail: v.m.dwyer@lboro.ac.uk

tion that cluster failures make to the observed early failure spectrum. In particular, we aim to calculate the probability of cluster failure P_{cf} for a given line (line length L, linewidth w, lognormal grain size distribution characterized by a median value of d_{50}, and a lognormal standard deviation or shape factor of σ_d) subject to a given set of test conditions (current density j and temperature T).

In the early 1980's Vaidya and Sinha noted the effects of the lognormal standard deviation σ_d on failure times in copper-doped aluminum interconnect, and proposed an empirical relationship between the mean time to failure ($\overline{t_f}$) and the microstructure of 0.5% Cu films, as[10]

$$\overline{t_f} \propto \frac{d_{50}}{\sigma_d^2} \ln\left(\frac{I_{111}}{I_{200}}\right)^3, \tag{1}$$

where $I_{(111)}$ and $I_{(200)}$ are the x-ray intensities of the (111) and (200) diffractions. Equation (1) predicts that a doubling of σ_d from 0.27 to 0.54 will reduce $\overline{t_f}$ by a factor of 4. Consequently, an additional area of interest will be the dependence of P_{cf} on σ_d. This is particularly important as simulated grain structures using the web-based grain growth programme MIT/EmSim (Refs. 6–8 and 11) possess a σ_d value of 0.27 while empirical, as-patterned grain structures possess a σ_d which appears to be closer to 0.54.[12,13]

The calculation in the current work is based on a one-dimensional model of the microstructure[14] which accurately predicts the simulated cluster-length distribution for $\sigma_d = 0.27$ but which also allows σ_d to be set as a free parameter. While clearly there are aspects of a two-dimensional grain structure which cannot be captured with a 1D representation, it may turn out that it is more important to have a realistic lognormal standard deviation σ_d.

The cluster failure probability P_{cf} is expected to increase with line length, linewidth/median grain size ratio (w/d_{50}) and with increasing variation in the grain size distribution σ_d. P_{cf} is expected to depend very weakly upon the diffusivity of the grains, with the proviso that the effective diffusion in clusters occurs at a substantially greater rate than in bamboo grains. The reason for this is that, although diffusion rates determine the times to failure, the critical lengths are determined by steady-state (or quasi-steady-state) stress profiles, which are roughly independent of the diffusivities.

A recent publication[14] set out a simple model for describing the microstructure of near bamboo lines based around the theory of runs.[15] The model generates a microstructure by choosing a grain size d from a lognormal distribution (d_{50}, σ_d) and assigning it as either a bamboo (or spanning) grain or a cluster grain according to its size relative to the linewidth w. A model of the line's microstructure is then reduced to a sequence of Bernoulli trials with success probability p (the probability that a grain is part of a cluster) equal to w/d. If all the grains are the same size d_0 then the probability that there exists a cluster longer than the critical length (say between M and $M+1$ grains), in a line of length N grains, is roughly $N(1-p)p^{M+1}$. Such rare structures are important as potential early failure sites.[5]

For a void to nucleate, the cluster must be long enough for it to develop the stress required for the nucleation [perhaps around 600 MPa (Ref. 16)]. The critical cluster length is expected to be related to the Blech length $L_B^{(n)}$ for nucleation in short via-via lines, although it will not be equal to $L_B^{(n)}$ as the bamboo regions at the ends of the cluster are not as efficient as tungsten vias in their role of blocking boundaries. Thus an interior cluster has its own critical length for nucleation ($L_C^{(n)}$), which is related to $L_B^{(n)}$ (e.g., Ref. 11), and is given by

$$L_C^{(n)} = \kappa L_B^{(n)} = \kappa \frac{2\sigma_{cr}\Omega}{Z^* q\rho j}, \tag{2}$$

where κ (of order 1.16 here) depends only upon the ratio of the effective diffusivities in the cluster and bamboo regions $1/\Gamma = \delta D_{GB}/wD_B$ [taken here as 80 (Ref. 16)], and the reduced critical stress $\Sigma_{cr} = \sigma_{cr}\Omega/kT$ (around 1.2, corresponding to $\sigma_{cr} = 500$ MPa). A complete list of parameters and their assumed values is given in Sec. II A.

After void nucleation, the cluster must also be long enough for the void to grow sufficiently large to cause failure. According to Refs. 4 and 5, and ignoring for the moment the influence of any initial thermal stress left from the processing stages, a critical void volume per cross section (V_C) for failure, assumed to be around 0.1–0.15 μm,[4,5,17] defines a critical cluster length for the void growth as[4]

$$L_C^{(g)} = \sqrt{\frac{2B\Omega V_C}{Z^* q\rho j}}. \tag{3}$$

The derivation of Eq. (3) in Ref. 4 assumes perfect blocking at the anode end of the cluster and at the cathode end of the void (i.e., $\delta D_{GB}/wD_B \to \infty$) and it is likely that a multiplicative constant λ (>1 and probably similar in magnitude to κ) should be inserted in here, as in Eq. (2), to account for the imperfection of the blocking by the bamboo grains at those boundaries.

The current density dependence of $L_C^{(g)}$ is different from that of L_C^n and it is clear that $L_C^{(n)} < L_C^{(g)}$ for sufficiently large current density. Consequently for high j we may expect that the growth of the void is the limiting factor—not all of the clusters of sufficient length to allow void nucleation will subsequently grow to failure size. Likewise, if j is sufficiently small, $L_C^{(n)} > L_C^{(g)}$ and, for those values of j, all voids that nucleate will eventually grow to failure. There is a crossover current density at which void-growth limited failure takes over from void-nucleation limited failure and this may be obtained by setting $L_C^{(n)} = L_C^{(g)}$. For a typical set of parameters, detailed below, and with $\kappa \approx \lambda$, the transition occurs at a current density of around $j_{trans} = 1.6$ MA cm^{-2} where $L_C^{(n)} = L_C^{(g)} = 12.5$ μm. This value is similar to that obtained by Park *et al.* for the $n=-2$ (nucleation) to $n=-1$ (growth) transition in the current density exponent of the median time to failure t_{50}.[17] Note that, because κ/λ is close to unity, this transition occurs for voids growing at the ends of polygranular clusters in long lines at roughly the same current density as for voids growing at the cathode via in short lines, and as a consequence, perhaps, in lines of all length. This result, if it is accurate, is remarkable in that, for all interconnect lines, under typical operational conditions $j < j_{trans}$ (i.e., the line operates in the region where nucleation is critical, in that if a

void nucleates in a polygranular segment the line will fail) while under typical test conditions $j > j_{trans}$ (i.e., the line operates in the region where growth is critical and voids may nucleate in polygranular segments without the line failing).

II. FAILURE MODEL

The model we shall use is based on both the theory of runs[14] and on the equation for evolution of tensile stress in a line described by Korhonen *et al.*[3] There are a number of variations to this equation;[17–19] however, the overall results in each case are similar.

A. Stress development

The model of Korhonen *et al.* determines the development of tensile stress within the interconnect according to the nonlinear continuity equation[3]

$$\frac{\partial \sigma}{\partial t} = \frac{\partial}{\partial x}\left[D_{eff} \exp\left(\frac{\sigma\Omega}{kT}\right)\left(\frac{\partial \sigma}{\partial x} - \frac{Z^* q\rho}{\Omega} j\right)\right] = -B\Omega\frac{\partial J}{\partial x}. \tag{4}$$

Equation (4) has the form of a nonlinear diffusion equation, however, the effective diffusivity D_{eff} is both averaged across the interconnect width, and so contains a factor δ/w for the case of grain boundary diffusion, and also contains the factor $B\Omega/kT$ ($\sim$120) which arises from the Hooke's law relationship between the stress and the atomic concentration. The model of Korhonen *et al.* is used here, rather than some of the derivative models,[17–19] partly for its ability to describe situations seen experimentally and partly because, despite its complexity, it is still relatively simple. In addition, amongst those 1D models which assume equilibrium statistics,[17–19] the predictions relevant to the current work are largely independent of the model used.

The equation for the stress development, Eq. (4), is solved using a standard finite-difference scheme, with normal (symmetric) central differences for first and second spatial derivatives, except at the line ends and at regions of discontinuity in the diffusivity, where a second order, asymmetric rule is used. The results are found to be independent of both the integration time step Δt and the step length Δx, and the scheme is perfectly stable against oscillations provided, as usual, that $D_{eff}(\sigma)\Delta t/\Delta x^2$ is kept within range. Verification calculations for the model are discussed elsewhere.[20]

The complete problem is defined according to the following parameter list.

(i) It is assumed that the critical tensile-stress defining void nucleation is 500 MPa, which is within the standard quoted range of 100 MPa–1 GPa and close to the deduced value of 600 MPa reported by Thompson *et al.*[16] In the present model short-circuit failure by dielectric rupture or extrusion has been ignored throughout. The atomic volume Ω is taken to be $1.66\times10^{-29} m^3$ (e.g., Ref. 16) and the temperature as 500 K.[16] A reduced (dimensionless) stress is defined as $\Sigma = \sigma(\Omega/\kappa T + 1/B) \approx \sigma\Omega/\kappa T$, which consequently has a critical value of $\Sigma_{cr} = 1.2$. The elastic modulus B is assumed to be 50 GPa, and κ is the Boltzmann's constant.

(ii) The critical void volume (per cross section) for failure V_C is taken to be 0.125 μm, which is within the range 0.1–0.15 μm quoted in the literature.[4,5,17] This value of the parameter V_C was deduced for 1 μm $\times$ 1 μm lines.[17] In order to use the theory of runs we shall assume the lines have submicron linewidths. However, we shall take this same critical value for the void length (volume/cross section). Effectively this assumes that the critical void volume scales with the line cross section. Equation (3) then provides a critical product for void growth $A_B^{(g)} = j(L_B^{(g)})^2$, given by $BV_C A_B^{(n)}/\sigma_{cr}$, of 2.5 A. At a current density of 2×10^6A cm^{-2}, this implies a critical length $L_B^{(n)}$ of around 1.1×10^{-3} cm, or 11 μm, for short line via-via failure.

(iii) The effective electromigration charge number is taken as $Z^* = 10$,[21] the resistivity for copper-doped aluminum is assumed to be 5×10^{-8} Ω m (Ref. 21) and q is the electronic charge. For short line, via-via failure this leads to a threshold product of $A_B^{(n)} = 2000$ A cm^{-1}, giving a critical length of $L_B^{(n)} = 10$ μm ($< L_C^{(g)} = 11$ μm) at a current density of 2 MA cm^{-2}, Note that at 1 MA cm^{-2}, $L_B^{(n)} = 20$ μm $> L_B^{(g)} = 16$ μm.

(iv) We also define the parameter $\beta = Z^*\rho q/\kappa T$ which, with these assumptions, and the current density again taken to be 2 MA cm^{-2}, gives the product $\beta j = 0.24$ μm^{-1}. For a characteristic length of, say, $\ell = L_B^{(n)} = 10.0$ μm, the product $\chi = \beta j \ell$ is then 2.4.

(v) The ratio between the effective diffusivity for the grain boundary diffusion and the bulk diffusion $\delta D_{GB}/wD_B$ is assumed to be fixed at around 80,[16] which is typical of quoted values.[16,5–9] The dependence of the effective atom diffusivity on stress is given by the relation[3]

$$D_A(\sigma) = D_A(0)\exp\left[\left(\frac{\Omega}{kT} + \frac{1}{B}\right)\sigma\right] \approx D_A(0)\exp\left(\frac{\Omega\sigma}{kT}\right). \tag{5}$$

(vi) The lines are pad-pad [or via-via (or stud-stud) with large metal reservoirs surrounding the studs/vias], of linewidth $w = 0.2$–0.6 μm and of length greater than 100 μm. Lines are assumed to be made from aluminum grains drawn from a lognormal distribution with a median grain size of 2 μm and a lognormal standard deviation of σ_d in the range 0.25–0.6.[7,12,13]

(vii) Thermal stress σ^T resulting from the manufacturing process is ignored, although its effects may be included in a straightforward manner and are discussed in Sec. II E.

If polygranular clusters exist within the line, where the effective diffusivity is $(B\Omega/kT)\times\delta D_{GB}/w$, the failure criterion can be met significantly more quickly provided that the cluster is sufficiently long. Indeed, as with via-via failure, a threshold product also operates in this case, although it will be different from $L_B^{(n)}$. This can be seen easily by considering an isolated polygranular cluster of length L_C. Here the line may be taken to be effectively infinite with the cluster placed centrally and with simple boundary conditions $\sigma = 0$ at the ends. These boundary conditions are chosen so that the ends do not influence the stress buildup in the cluster. The line is thus divided into the union $x = (-\infty, 0)$

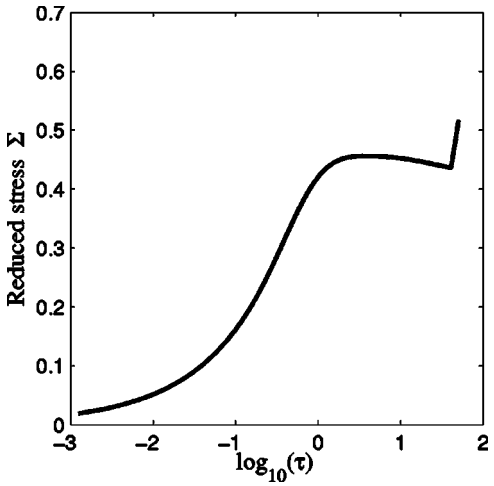

FIG. 2. The maximum stress in the line as a function of reduced time (τ) on a $\log_{10}$ scale. The interconnect here consists of a line of length 100 μm with a 20 μm cluster placed 10 μm from the cathode. The current density j in this case is 2 MA cm^{-2}.

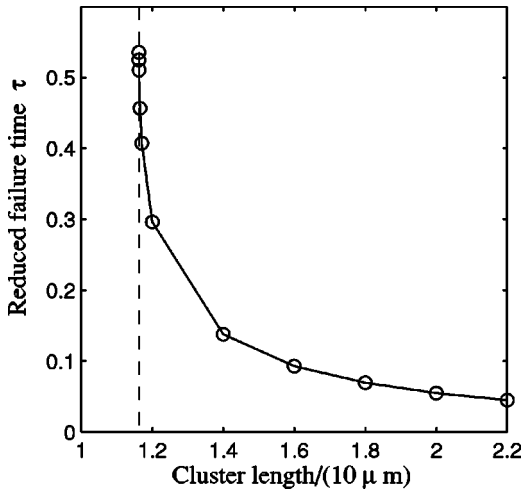

FIG. 3. Reduced failure time (τ) as a function of cluster length in units of 10 μm which is the critical (Blech) length for via-via failure. The reduced time becomes effectively infinite when the cluster length $L_c^{(n)}$ falls below around 1.162, or 11.62 μm when the failure is slow. The current density is $j=2$ MA cm^{-2}.

$\cup[0,L_C]\cup(L_C,\infty)$ where fast diffusion takes place only in the central region $x=[0,L_C]$. Introducing the following dimensionless units:

$$\Sigma=\frac{\Omega\sigma}{kT},\quad X=\beta jx,\quad \tau=(\beta j)^2\frac{B\Omega}{kT}\frac{\delta D_{GB}}{w}t,$$

$$\frac{1}{\Gamma_{eff}}=\frac{\delta D_{GB}}{wD_{eff}}\frac{B\Omega}{kT}, \tag{6}$$

the stress equation may be written as

$$\frac{\partial\Sigma}{\partial\tau}=\Gamma_{eff}\frac{\partial}{\partial X}\left[\exp(\Sigma)\left(\frac{\partial\Sigma}{\partial X}-1\right)\right] \tag{7}$$

on the set $X=(-\infty,0)\cup[0,X_C=\beta jL_C]\cup(X_C,\infty)$, with failure criterion $\Sigma=\Sigma_{cr}=\Omega\sigma_{cr}/kT$ and end conditions $\Sigma=0$. The parameter $\Gamma_{eff}=1$ in the cluster regions and $\Gamma_{eff}=\Gamma=wD_B/\delta D_{GB}(=1/80)$ in the bamboo regions. Notice that in the determination of the reduced stress $\Sigma(X,\tau)$, the current density now only appears in the reduced length $X_C=\beta jL_C$, and consequently once the product jL_C is fixed, X_C is fixed and the problem becomes well defined. As a consequence jL_C determines the solution, and any critical behavior corresponds to a critical value of this product. For a given value of jL_C we may determine a unique stress profile and a unique reduced fast failure time τ_f, which only depend upon the reduced parameters Γ and Σ_{cr}.

A typical time variation for the maximum tensile stress which develops within the line which contains a cluster, $\Sigma_{\max}(\tau)=\max_X\Sigma(X,\tau)$, is shown in Fig. 2. There is an early increase in $\Sigma_{\max}(\tau)$ due to atom migration in the cluster until a quasi-steady-state is reached on a time scale of order $(L_B^{(n)})^2/D_{eff}^{(GB)}$. This is followed by a slower increase requiring transgranular (or other slow) diffusion on a time scale of order $(L_B^{(n)})^2/D_{eff}^{(B)}$ and finally a gradual reduction which arises because of the nonlinear nature of the problem and is not seen in any linearized versions, such as Refs. 3 and 17, but is apparent, for example, in Fig. 6 of Ref. 6. Clearly, there is a critical point $X_C=X_C^*$ at which the reduced cluster length becomes too low to support failure—this will be the greatest reduced length X_C at which the maximum value of the maximum stress profile $\max_{X,\tau}\{\Sigma(X,\tau)\}$ is less than Σ_{cr}. For clusters below the critical length the detailed conditions at the line boundaries become important and it is these that will eventually determine the nucleation time. However, on the time scale of interest this is effectively a $\tau_f\to\infty$ (and hence $t_f\to\infty$) transition. Failure cannot now be termed early or cluster failure and therefore X_C^* determines a critical value for nucleation for the product $jL_C^{(n)}$.

Figure 3 shows the failure time as a function of cluster length using the assumed values above. Here we have taken the unit of distance to be 10 μm, which is the via-via Blech length for $j=2$ MA cm^{-2}. Unlike $jL_B^{(n)}$, which only depends upon the critical stress, Eq. (1), the value for $jL_C^{(n)}$ also depends upon the ratio of the effective grain boundary and the effective bamboo diffusion. Thus

$$(jL_C^{(n)})_{cr}=\frac{X_C^*(\Sigma_{cr},\Gamma)}{\beta}=\frac{2\sigma_{cr}\Omega}{Z^*e\rho}\frac{X_C^*(\Sigma_{cr},\Gamma)}{2\Sigma_{cr}}=\frac{X_C^*(\Sigma_{cr},\Gamma)}{2\Sigma_{cr}}(jL_B^{(n)})_{cr}. \tag{8}$$

The constant of proportionality $X_C^*/2\Sigma_{cr}$ depends upon the ratio of effective diffusivities in the cluster and bamboo regions (around 80 here), and also the reduced critical stress Σ_{cr} (around 1.2). Its value is around 1.16, leading to a threshold product of $jL_C^{(n)}=2$ MA cm$^{-2}\times 11.6$ μm$=2320$ A cm^{-1}, which compares to 2000 A cm^{-1} for short lines blocked by tungsten vias.

If the critical value for void nucleation is exceeded, a void will form at the cathode end of the cluster and it will start to grow. The stress at the edge of the void collapses to zero, due to the presence of the free surface. If the initial void volume after nucleation is Δ_0, integrating the stress equation (4) along the cluster length L_C yields

$$\frac{1}{B\Omega}\frac{\partial}{\partial t}\int_0^{L_C}\sigma\, dx = J(0,t) - J_{void}(t) \approx -J_{void}(t), \tag{9}$$

where we have assumed that the cluster is effectively blocked at its anode-most end by bamboo grains and thus $J(0,t)\approx 0$. Integrating Eq. (9) between the void nucleation time t_{nucl} and time t gives

$$\frac{1}{B\Omega}\int_0^{L_C}\sigma(x,t)dx = \frac{1}{B\Omega}\int_0^{L_C}\sigma(x,t_{nucl})dx - \int_{t_{nucl}}^{t} J_{void}(t)dt. \tag{10}$$

Up to the point of nucleation the cluster effectively acts as though it is blocked at both ends. Consequently, the first integral on the right-hand side is close to zero. Assuming that vacancies are also blocked at the cathode end of the void, the increase in void size at time t is $\Omega hw\int_{t_{nucl}}^{t} J_{void}(t)dt$. Thus the void volume at time t is

$$V(t) = \Delta_0 - hw\int_0^{L_C}\frac{\sigma(x,t)}{B}dx, \tag{11}$$

where Δ_0 is the initial void size. As $t\rightarrow\infty$ the stress profile roughly becomes $\sigma(x,t\rightarrow\infty) = -Z^*q\rho j(L_C - x)/\Omega$, so that the maximum void volume is

$$V_{\max} = \Delta_0 + hw\frac{Z^*q\rho jL_C^2}{2B\Omega}. \tag{12}$$

Setting $V_C = (V_{\max} - \Delta_0)/hw$, one regains Eq. (3). The development of Eqs. (9)–(12) is more or less identical to that given by Korhonen *et al.* in Ref. 3. It is repeated here because the same argument can be extended to several of the other equilibrium, 1D models used for electromigration modeling,[17–19] and to make clear that the time integrated electromigration current density $J_{void}(t)$ into the void is simply related to the stress profile integrated over the entire cluster length [Eq. (10))]. This has been leveled as a criticism[17] of the model of Korhonen *et al.*, but in fact arises as a result of the effective blocking boundary at the anode end of the cluster and at the cathode end of the void so that any increase/decrease in the total compressive stress occurs due to the current of vacancies into/out of the void. These atoms can naturally be incorporated anywhere in the cluster. This argument is independent of the electromigration models considered,[3,17–19] although naturally the details are somewhat model dependent.

The validity of this argument may be seen in the following manner. First, all the equilibrium-1D electromigration models under consideration[3,17–19] assume the same expression for the vacancy flux and thus the same quasi-steady-state ($J=0$) solution, in which the stress in the interconnect is a linear function of displacement x from the anode end of the cluster. Second, all may be written in the form of a continuity relation. Thus

$$h(\sigma)\frac{\partial\sigma}{\partial t} = -\frac{\partial J}{\partial x} \quad \text{or} \quad \frac{\partial H(\sigma)}{\partial t} = -\frac{\partial J}{\partial x}, \tag{13}$$

where $h(\sigma) = H'(\sigma)$ is a simple function of stress. In the model of Korhonen *et al.*,[3]

$$H(\sigma) = C_0\left[1 - \exp\left(-\frac{\sigma}{B}\right)\right] \approx \frac{\sigma}{B\Omega}, \tag{14}$$

while in Clement's model,[18] where atomic recombination is assumed to occur only in grain boundaries,

$$H(\sigma) = C_0\left[1 - \exp\left(-\frac{\sigma}{B}\right)\right] + \frac{\delta}{w}C_{v0}\left[\exp\left(\frac{\Omega\sigma}{kT}\right) - 1\right] \tag{15}$$

and in the model of Park *et al.*,[17] where atomic recombination occurs equally through out the cross section but vacancy formation requires an energy Ev,

$$H(\sigma) = C_0\left[1 - \exp\left(-\frac{\sigma}{B}\right)\right] + C_{v0}\left[\exp\left(\frac{\Omega\sigma}{kT}\right) - 1\right] \times \exp\left(-\frac{E_V}{kT}\right). \tag{16}$$

Here C_0 is the atomic concentration under zero stress ($=\Omega^{-1}$), C_{v0} is the vacancy concentration under zero stress. Indeed, even the original electromigration model of Shatkzes and Lloyd[19] may be cast in this simple form by setting

$$H(\sigma) = C_{v0}\left[\exp\left(\frac{\Omega\sigma}{kT}\right) - 1\right]. \tag{17}$$

It is clear that all models represented by Eqs. (14)–(17) are special cases of a more general expression involving both vacancy terms as in Eq. (17) and lattice terms as in Eq. (14).

As a consequence of Eq. (13), whichever of these models we choose

$$\frac{\partial}{\partial t}\int_0^{L_C}H(\sigma)dx = J(0,t) - J_{void}(t) \approx -J_{void}(t) \tag{9'}$$

assuming again that the anode end of the cluster is effectively blocked, i.e., $J(0,t)$ is zero. Integrating Eq. (9′) over t as in Eq. (10), and substituting $H=0$ at $t=0$ (no initial thermal stress term) and $H=H(\sigma(x,t\rightarrow\infty))$ at $t\rightarrow\infty$, we arrive at a maximum void volume of

$$V_{\max} = \Delta_0 + hw\frac{\Omega}{kT}\int_0^{L_C} -H\left(-\frac{Z^*q\rho j}{\Omega}(L_C - x)\right)dx. \tag{12'}$$

The integral in Eq. (12′) is straightforward for the cases (14)–(17), each of which generate a simple relationship between the assumed critical void volume and the related critical cluster length. Equation (12′) shows that the current of vacancies pouring into/out of the void is represented in each case by an integral over the entire stress profile and this does not separate the models governed by Eqs. (14) and (16) as has been suggested in Ref. 17. In each case, expanding the power series for the exponential terms appearing in Eq. (12′) yields, to lowest order in L_C, relationships between $V_{\max}$ and

TABLE I. Exact values and approximations to P_{cf}, for Blech lengths of (a) 15 μm and (b) 7 μm. At 15 μm 1% failure is not reached until the line length is 10 mm. The integer in the bracket is an exponent to base 10. Here z_1 and z_2 correspond to the left and right inequalities in Eq. (18). Exact results use the embedded Markov method illustrated in Ref. 24.

	(a) 15 μm			(b) 7 μm		
Line length (μm)	Exact	$1-z_1$	$1-z_2$	Exact	$1-z_1$	$1-z_2$
40	2.713 6(−5)	3.327 9(−5)	2.662 4(−5)	0.0219	0.0269	0.0219
100	8.857 4(−5)	11.007(−5)	8.806 0(−5)	0.0590	0.0725	0.0588
200	1.909 6(−4)	2.380 5(−4)	1.904 5(−4)	0.1177	0.1439	0.1173
2000	0.002 032	0.002 539	0.002 032	0.7230	0.7974	0.7218

L_C, all of which are of the form $V_{\max}=\Delta_0+aL_C^2/2$ and thus all of which produce a version of Eq. (4).

Equations (8) and (12) or (12′) define the critical lengths for the cluster. However, to obtain the probability of such structures, for the nucleation and subsequent growth to failure of voids at interior points on an interconnect, we also need a means of modeling the microstructure. We do this here using the model developed in Ref. 14.

B. Theory of runs

In Ref. 14 we demonstrated that the microstructure of fine lines could be described in terms of a simple one-dimensional model. Within this model the grain size distribution was assumed to be lognormal, and near bamboo lines were created by first assigning the grain sizes d and then determining the grain status (bamboo B, or cluster grain G) as a sequence of Bernoulli trials with a success probability (probability of the grain having a boundary lying along the interconnect) of w/d. It was demonstrated in Ref. 14 that this simple model is able to reproduce, more or less exactly, the cluster length distribution and its variation on linewidth generated by the 2D simulator MIT/EmSim.[6–8] Additionally, the lognormal standard deviation of the grain size distribution enters as a free parameter, whereas for the 2D simulator it appears as a result of the growth process.

C. The $\sigma_d=0$ model

It is useful to consider first the simple, but important, case in which the interconnect contains grains of the same size. A line of length L is assumed to correspond to a set of $N=L/d_0$ grains, each of diameter d_0, which are cluster grains (G) with probability $p=w/d_0$ and bamboo grains (B) with probability $q=1-p$. Each possible interconnect microstructure can then be uniquely specified by a codeword generated by a Bernoulli trial. For example, *BBGGGGBBBGGBBGGB* would correspond to an interconnect of length $L=16\ d_0$ (or around 32 μm) containing one cluster of length $4d_0$ (i.e., $L_{\max}=4d_0$) and two of length $2d_0$. The smaller clusters are unimportant for the early failure probability, what is important is whether or not $L_{\max}=4d_0>L_{crit}=\max(L_C^{(n)},L_C^{(g)})$, the critical length for the particular operating conditions. The presence of clusters may then be calculated using the theory of runs from standard probability theory.[22,23]

As an example, if the linewidth is taken as 0.4 μm and, with a grain diameter $d_0=2$ μm, the probability of "success" $p=w/d_0=0.2$. With this model, the exact probability[24] that the longest connected cluster is greater than $3d_0=6$ μm and $7d_0=14$ μm (corresponding to critical cluster lengths of, for example, 7 μm and 15 μm, respectively) may be calculated using the method of Markov embedding and is shown (under the heading "Exact") for a variety of stripe lengths in Table I. In a 2000 μm test structure the probability of finding a cluster of length greater than 15 μm is 0.2% while in lines of 200 μm, the failure probability is 0.02%. If the current density is so high that the critical length lies in the interval 6 μm $<L_{crit}<8$ μm (using $jL_{crit}=2000$ A/cm,[25] this requires j above 2.5 MA cm^{-2}), the probability of such a failure in 200 μm lines rises to around one in nine. Approximate results, which may be more easily adapted to the case of $\sigma_d\neq 0$, can be obtained in the following manner.

We suppose that, in general, the critical length lies in the range $[Md_0,(M+1)d_0]$ so that failure will occur if the longest cluster is greater than Md_0. Define $P(m)$ as the probability that a cluster of m grains has a length greater than the critical length, and thus constitutes a failure unit. For $\sigma_d=0$, this is clearly zero if $m\leqslant M$ and unity if $m\geqslant M+1$, i.e., $P(m)=\Theta(m-M-1/2)$, where $\Theta(m)$ is the unit step function. For finite σ_d, $P(m)$ will not be a simple step function but something more gradual, Sec. II D.

The probability that there are k clusters of length m in a line of N trials (grains) is approximately[22]

$$p_k(m)=\frac{(Nqp^m)^k}{k!}\exp(-Nqp^m). \tag{18}$$

The probability that the maximum cluster length is less than L_{crit} is the probability that, for all $m>0$, all $k(m)$ clusters of length m are less than L_{crit}. This is the product of the probabilities, for all m, that all $k(m)$ clusters of length m are less than L_{crit}. That is,

$$\Pr\{L_{\max}\leqslant L_{crit}\}=\prod_{m=1}^{\infty}\sum_{k=0}^{\infty}p_{k(m)}(1-P(m))^{k(m)}$$
$$=\exp\left(-\sum_{m=M+1}^{\infty}Nqp^m\right)=\exp(-Np^{M+1}). \tag{19}$$

Consequently, the probability of early failure P_{cf} is equal to $=1-\exp(Np^{M+1})$.

We can justify the approximation in Eq. (18) and the resulting Eq. (19) in the following manner. It is possible to show rigorously that (Ref. 24 and references therein)

$$\exp(-Np^{M+1}) \approx (1-p^{M+1})^{N-M} \leqslant P(L_C < (M+1)d_0) \leqslant (1-qp^{M+1})^{N-M} \approx \exp(-Nqp^{M+1}). \tag{20}$$

Notice that in the present case where M is of order of perhaps 4–12 grains, and $p=0.1-0.6$, the difference between the upper and lower bounds is very small so that we obtain a very accurate estimate of $P(L_C<(M+1)d_0)$. This accuracy is demonstrated in Table I where z_1 and z_2, corresponding to the left- and right-hand inequalities in Eq. (20), are compared to the exact values obtained by the method of Markov embedding.[24]

D. Lognormal grain size distribution ($\sigma_d \neq 0$)

The grain size distribution has been found to follow a fairly tight lognormal distribution with quoted values of the shape factor (lognormal standard deviation) σ_d of around 0.4–0.6.[12,13] The 2D simulator MIT/EmSim generates realistic grain structures but leads to a shape factor of $\sigma_d=0.27$.[6–8] Consequently our range of values for σ_d is chosen to be 0.2–0.6. The effect of a nonzero shape factor is to smear the step nature of $P(m)$. One might expect most of the conclusions above to remain true, however.

It is possible[14] to separate the trials from the choice of grain sizes by changing the grain size distribution to a lognormal distribution with a median value of $d'_{50}=d_{50}\exp(\sigma_d^2)$ rather than the original distribution, and by setting $p=(w/d_{50})\exp(\sigma_d^2/2)$. Essentially, local variations in the values of $p=w/d$, due to variations in grain size, can instead be included in the cluster length distribution by altering the success probability p and the median grain size. For analysis it is also convenient to approximate the lognormal grain size distribution by a Γ distribution with location $a(\sigma_d)$, shape factor $b=3$, and scale factor $c(\sigma_d)$.[14] The probability $\Pr\{S_m<L_{crit}\}$ that a cluster of m grains is less than the critical cluster length is then related to the incomplete Gamma function[14] $\Gamma(3m,x=(L_{crit}-a)/c)/\Gamma(3m)$. Note, for $L_{crit}<ma, P(S_m<L_{crit})=0$ and for $L_{crit}>ma$,

$$P(m)=\Pr\{S_m>L_{crit}\}=1-\int_{ma}^{L_{crit}}\frac{(x-ma)^{3m-1}}{c^{3m}\Gamma(3m)}\times\exp\left(-\frac{x-ma}{c}\right)dx = e_{3m-1}([L_{crit}-ma]/c)\exp(-[L_{crit}-ma]/c), \tag{21}$$

where $e_n(x)$ represents the sum of the terms, up to the term in x^n, in the power series expansion of $\exp(x)$.[14] The first equality in Eq. (19) still describes the probability that the maximum cluster is less than L_{crit}, provided that the appropriate value of $P(m)$ is used. Thus we obtain the final result for P_{cf} as

$$1-P_{cf}=\Pr\{L_{\max}\leqslant L_{crit}\}=\prod_{m=1}^{\infty}\sum_{k=0}^{\infty}p_{k(m)}(1-P(m))^{k(m)} = \exp\left(-\sum_{m=1}^{\infty}Nqp^mP(m)\right), \tag{22}$$

where $P(m)$ given in Eq. (21). Note that in the limit of $\sigma_d \to 0$, where $P(m)=\Theta(m-M-1/2)$, and Eq. (19) is regained.

E. Effects of initial thermal stress

The initial stress $\sigma^{\mathrm{T}}(x)$ from the manufacturing process has been neglected to this point; however, it is a simple matter to include such stress and its consequences can be significant. From the point of view of void nucleation, $\sigma^{\mathrm{T}}(x)$ reduces the stress required from the metal migration process, leading to an effective critical stress at x of $\sigma_{cr}-\sigma^{\mathrm{T}}(x)$. The critical length for void nucleation close to x, $L_C^{(n)}$, will be reduced from Eq. (2) by a factor of $1-\sigma^{\mathrm{T}}(x)/\sigma_{cr}$ as it is proportional to the required stress from metal migration. When the value of $L_C^{(n)}$ dominates the failures process, an initial thermal stress of $\sigma_{cr}/2$ will typically reduce the critical length from M grains to $M/2$ grains.

The critical cluster length condition for void growth with initial stress ($L_C'^{(g)}$, say), has been already obtained by Korhonen *et al.*[5] (it also may be obtained from Eq. (10) by evaluating the first integral on the right-hand side to be $\sigma^{\mathrm{T}}L_C'^{(\mathrm{g})}$ rather than zero) as

$$\frac{Z^*q\rho j(L_C'^{(g)})^2}{2kT}+\frac{\Omega\sigma^T}{kT}L_C'^{(g)}-\frac{B\Omega}{kT}\frac{(V_{cr}-\Delta_0)}{hw}=0. \tag{23}$$

This may be written for the fractional change in $L_C'^{(g)}$ compared to case of the zero initial stress (i.e., $z=L_C'^{(\mathrm{g})}/L_C^{(g)}$) as

$$z^2+2\delta z-1=0, \tag{24}$$

where $\delta=\sigma^TL_C^{(n)}/2\sigma_{cr}L_C^{(g)}$. At current densities for which $L_C^{(g)}\sim L_C^{(n)}$, and with an initial stress $\sigma^{\mathrm{T}}<\sigma_{cr}/2$, δ is relatively small and we may approximate $L_C'^{(g)}\approx(1-\sigma^{\mathrm{T}}/2\sigma_{cr})L_C^{(g)}$. Thus the effect of initial stress here is about half as great as on $L_C^{(n)}$ but can still be significant. At low current densities $\sim$0.1 MA cm^{-2}, where $L_C^{(n)}\approx 200$ μm and $L_C^{(g)}\approx 50$ μm, the ratio $L_C^{(n)}/L_C^{(g)}\approx 4$. If $\sigma^{\mathrm{T}}=375$ MPa (i.e., $0.75\sigma_{\mathrm{cr}}$), $\delta\sim 1.5$ and $L_C'^{(g)}=0.3L_C^{(g)}\approx 15$ μm. While a cluster of 50 μm represents a highly unlikely event in 100 μm lines with $p=0.2$, a cluster of 15 μm cluster will occur with a probability of 8.9×10^{-5} in lines of length 100 μm, Table I, when the shape factor $\sigma_d=0$. As always, this value will increase with σ_d.

It is clear that the early failure problem becomes more severe in the presence of high initial stress. For the parameters listed in Sec. II A, and with a current density of $j=2$ MA cm^{-2}, the critical cluster lengths are $L_C^{(g)}=12.8$ μm and $L_C^{(n)}=11.6$ μm. If $\sigma^T=\sigma_{cr}/2=250$ MPa, $L_C'^{(n)}=5.6$ μm and $L_C'^{(g)}=10.2$ μm and the critical cluster length is thus reduced from 12.8 μm to 10.2 μm (both growth dominated). This reduction is around one grain size, so that for a line-

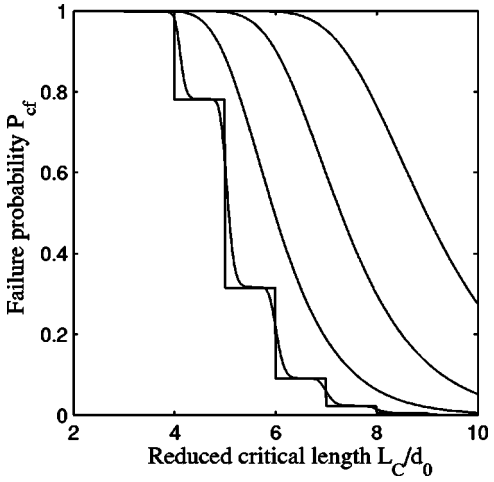

FIG. 4. Failure probability P_{cf} as a function of reduced cluster length (units of d_{50}) for σ_d values of 0 (leftmost), 0.05, 0.27 (the simulator value), 0.40 and 0.50 (rightmost). Here the line length $L=1550d_{50}$ and $p=0.25$ where $p=(w/d_{50})\exp(\sigma_d^2/2)$.

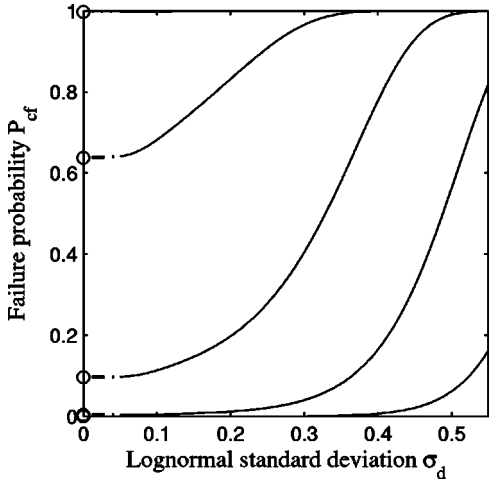

FIG. 5. Failure probability P_{cf} as a function of the lognormal standard deviation. The exact results for $\sigma_d=0$ are shown as circles. Plots for $L=7.5\ d_{50}$, around 15 μm and line length=1550 d_{50}. Plots are, from top to bottom, $p=0.5$ to 0.1 in steps of 0.1. At $p=0.5$ for this length failure is almost certain. As usual, $p=(w/d_{50})\exp(\sigma_d^2/2)$.

width of 0.5 μm, with $p=0.25$, the probability of a failing cluster is increased by around a factor of $1/p=4$.

Solving Eqs. (2) and (23) together for the transition current density between growth-dominated and nucleation-dominated failure, with a nonzero σ^{T} (and $\kappa=\lambda=1$), yields a value of $L_C=L_C^{(n)}=L_C^{(n)}$ which is independent of σ^{T} (12.5 μm). Thus the transition current density j_{trans} is also reduced by a factor $1-\sigma^{\mathrm{T}}(x)/\sigma_{cr}$.

III. RESULTS AND DISCUSSION

The current model [Eqs. (21) and (22)] allows us to assess the importance of the lognormal standard deviation (or shape factor σ_d) of the grain size distribution on the probability of early failure P_{cf} in near-bamboo as-patterned aluminum lines. It is clear from Fig. 4, that the value of P_{cf} is strongly dependent on the shape factor. This raises an important issue as currently quoted values of σ_d for real aluminum films are in the range 0.4–0.6 (Refs. 12, 13, and 26) while the simulator MIT/EmSim generates lognormal distributions with a shape factor of 0.27. Clearly, for values of the critical length around $6d_{50}$ or above (i.e., 12 μm and above in a line of width $w=pd_{50}=0.5$ μm) the failure probability for $\sigma_d=0.27$ is orders of magnitude better than for $\sigma_d=0.5$. The reason for this is that, although the simulator produces lines with realistic microstructures, the variance of the lognormal curve (for a fixed d_{50}) increases fairly rapidly with σ_d. For $\sigma_d=0.27$ the standard deviation is $0.285d_{50}$, while at $\sigma_d=0.5$ the standard deviation is $0.604d_{50}$ and, consequently, for a larger σ_d, the high probability clusters, those with a small number of grains, have a greater chance of causing failure.

It must also be remembered that in an IC with several million interconnects it is those clusters several standard deviations away from the cluster mean that will determine the failure probability. As a consequence, it is important to have an accurate value for the lognormal standard deviation.

Notice the convergence of the $\sigma_d=0.05$ curve to the $\sigma_d=0$ curve. Convergence is fastest if the critical cluster length is, as here, $(M+\frac{1}{2})d_{50}$. In Fig. 5, which shows the failure probability as a function of σ_d for a critical length of 15 μm (i.e., $M=7$) $L=1550d_0$ and p values of 0.1–0.5, this is seen as a flattening in the curves as $\sigma_d=0$ is approached. Were the critical length elsewhere in the interval $[Md_{50},(M+1)d_{50}]$ the flattening in the curves in Fig. 5 would still occur, but at a much smaller value of σ_d. Note in Fig. 4 that if $L_{crit}=13$ μm, i.e., $M=6$ the probability of failure for $\sigma_d=0.27$ is around 0.25 while for $\sigma_d=0.5$ it is around 0.97. In the important case where P_{cf} is small the result of the simulator may be an order of magnitude or more too optimistic.

Figure 6 shows the failure probability P_{cf} as a function of the linewidth w for a variety of values of σ_d and j. Figures 7 and 8 demonstrate the current density dependence of P_{cf} for, respectively, fixed σ_d (0 and 0.4) and a variety of p values (0.1–0.5) and fixed $p=0.25$ and a variety if σ_d values.

In conclusion, microstructural details are vital to line

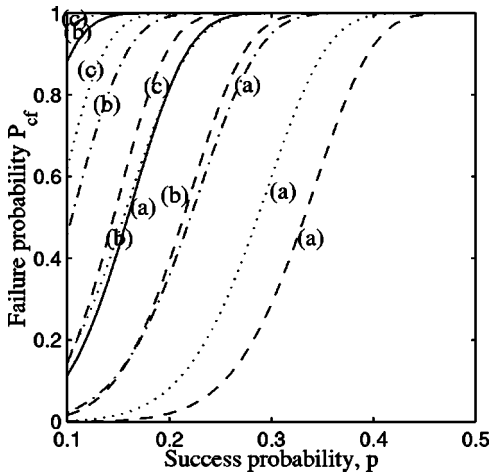

FIG. 6. Plots of the failure probability P_{cf} against the linewidth parameter $p=(w/d_{50})\exp(\sigma_d^2/2)$ for a range of parameters (dashed curves correspond to $\sigma_d=0$, dotted curves to $\sigma_d=0.27$, dash-dot curves to $\sigma_d=0.4$, and solid curves to $\sigma_d=0.5$). Here the current values in MA cm^{-2} are, respectively, (a) 1.5, (b) 2.25, and (c) 3. Line length=1550 d_{50}.

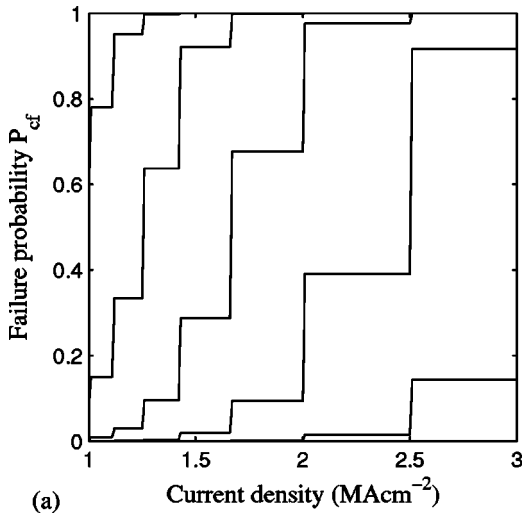

FIG. 7. Failure probability as a function of current density j in MA cm^{-2}. (a) corresponds to $\sigma_d=0$ and (b) corresponds to $\sigma_d=0.4$, p values of 0.1 to 0.5 inclusive, starting from bottom. The failure probability in (b) for $p=0.5$ is effectively unity. Line length$=1550\ d_{50}$. As usual, $p=(w/d_{50})\exp(\sigma_d^2/2)$.

failure and, in particular, an accurate model of it is required to assess the failure probability P_{cf} for early failure. It is probably true that this is the case for any failure mode and not merely fast failure. This means that it is crucial to get the grain size distribution right and as scaling means that d_{50} may be accurately set, and as the distribution follows a log-normal plot fairly tightly, this means that it is important to get σ_d correct. We propose a 1D model of the microstructure which accurately predicts the cluster distribution produced by the grain growth simulator MIT/EmSim Refs. 6–8 at $\sigma_d=0.27$, but which also allows σ_d to be set arbitrarily. The results show that, although undoubtedly something will be lost in reducing from a 2D model to a 1D model, for applications such as this, it is probably more important to use an accurate value for σ_d.

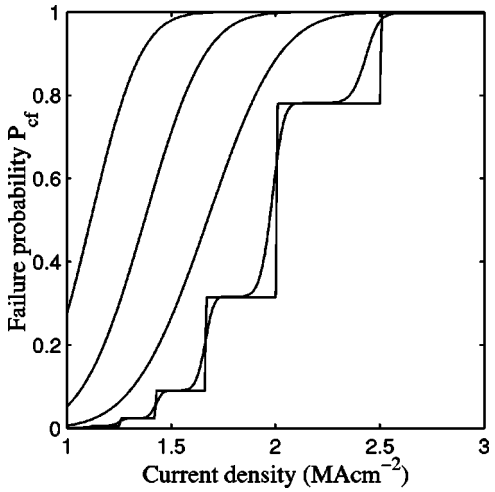

FIG. 8. Failure probability as a function of current density for σ_d values of 0 (rightmost), 0.05 and 0.27 (the simulator value), 0.40 and 0.50 (leftmost). Here the line length$=1550\ d_{50}$ and $p=(w/d_{50})\exp(\sigma_d^2/2)=0.25$.

Specifically, our analysis shows that P_{cf} is strongly dependent on σ_d, and a change in σ_d from 0.27 to 0.5 can cause an order of magnitude (or more) increase in P_{cf} for typical test conditions. The implications for the web-based 2D grain-growth simulator MIT/EmSim, which generates grain patterns with a σ_d value of 0.27 rather than the observed value in as-patterned structures of 0.4–0.6, are that the simulator is likely to significantly overestimate interconnect reliability due to this particular electromigration failure mode.

[1] I. A. Blech, J. Appl. Phys. **47**, 1203 (1976).
[2] I. A. Blech and C. Herring, Appl. Phys. Lett. **29**, 131 (1976).
[3] M. A. Korhonen, P. Børgesen, K. N. Tu, and C.-Y. Li, J. Appl. Phys. **73**, 3790 (1993).
[4] R. G. Filippi *et al.*, J. Appl. Phys. **91**, 5787 (2002).
[5] M. A. Korhonen, P. Børgesen, D. B. Brown, and C.-Y. Li, J. Appl. Phys. **74**, 4995 (1993).
[6] B. D. Knowlton, J. J. Clement, and C. V. Thompson, J. Appl. Phys. **81**, 6073 (1997).
[7] H. J. Frost, C. V. Thompson, and D. T. Walton, Acta Metall. Mater. **38**, 1455 (1990).
[8] B. D. Knowlton and C. V. Thompson, J. Mater. Res. **13**, 1164 (1998).
[9] Y. Liu, C. L. Cox, and R. J. Diefendorf, J. Appl. Phys. **83**, 3600 (1998).
[10] S. Vaidya and A. K. Sinha, Thin Solid Films **75**, 253 (1981).
[11] Y.-C. Joo and C. V. Thompson, J. Appl. Phys. **76**, 7342 (1994).
[12] K. Wu, W. Baerg, and P. Jupites, Appl. Phys. Lett. **58**, 1299 (1991).
[13] C. V. Thompson, Annu. Rev. Mater. Sci. **20**, 245 (1990).
[14] V. M. Dwyer, J. Phys. D **37**, 422 (2004).
[15] W. Feller, *An Introduction to Probability Theory and its Applications*, 2nd ed. (Wiley, New York, 1971), Vol. 1.
[16] C. S. Hau-Riege and C. V. Thompson, J. Appl. Phys. **87**, 8467 (2000).
[17] Y.-J. Park, V. Andleigh, and C. V. Thompson, J. Appl. Phys. **85**, 3546 (1999).
[18] J. J. Clement, J. Appl. Phys. **82**, 5991 (1987).
[19] M. Shatkzes and J. R. Lloyd, J. Appl. Phys. **59**, 3890 (1986).
[20] V. M. Dwyer (unpublished).
[21] H. V. Nguyen, C. Salm, F. G. Kuper, and A. J. Mouthaan, *Proc. SAFE/IEEE workshop*, November (2000).
[22] R. von Mises, *Mathematical Theory of Probability and Statistics*, edited by H. Geiringer (Academic, New York, 1964).
[23] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. (Wiley, New York, 1971), Vol. I.
[24] J. C. Fu, L. Wang, and W. Y. W. Lou, J. Appl. Probab. **40**, 346 (2003).
[25] J. R. Lloyd, Semicond. Sci. Technol. **12**, 1177 (1997).
[26] J. E. Palmer, C. V. Thompson, and H. I. Smith, J. Appl. Phys. **62**, 2492 (1987).