RC23679 (W0507-212) July 29, 2005
Mathematics

IBM Research Report

Minimum Void Size and 3-Parameter Lognormal Distribution for EM Failures in Cu Interconnects

Baozhen Li[1], Cathryn Christiansen[1], Jason Gill[1], Ronald Filippi[2],
Timothy Sullivan[1], Emmanuel Yashchin[3]

[1]IBM Systems and Technology Group
Essex Junction, VT 05452

[2]IBM Systems and Technology Group
Hopewell Junction, NY 12533

[3]IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

MINIMUM VOID SIZE AND 3-PARAMETER LOGNORMAL DISTRIBUTION FOR EM FAILURES IN CU INTERCONNECTS

Baozhen Li[1], Cathryn Christiansen[1], Jason Gill[1],
Ronald Filippi[2], Timothy Sullivan[1] and Emmanuel Yashchin[3]
IBM Systems and Technology Group, 1) Essex Junction, VT 05452; 2) Hopewell Junction, NY 12533
3) IBM Research Division, Yorktown Heights, NY 10598
(802) 769-7962; Fax: (802) 769-4287; e-mail: Lib@us.ibm.com

ABSTRACT

Broad failure time distributions were observed for line depletion electromigration in Cu interconnects for various structures without sufficient liner contact and via redundancy. The root cause for this behavior was identified as the sensitivity of failure times to the void size, shape and location. Application of traditional 2-parameter lognormal distribution model to corresponding stress data often results in very pessimistic EM lifetime projections. A 3-parameter lognormal distribution was found not only to fit the experimental data better, especially for the early portion of the failure time distributions, but also to generate more accurate lifetime projections for void-size-limited EM. Given the nature of EM wear-out, deeper consideration indicates that a 3-parameter lognormal distribution has a sounder physical basis than a 2-parameter lognormal distribution. The new parameter introduced in the model, the minimum failure time (X_0), scales with via size over several technology generations, further validating the minimum void size explanation. [*KEY WORDS:* Electromigration, Cu interconnects, voids, reliability, lognormal, redundancy.]

INTRODUCTION

Electromigration (EM) is one of the major reliability concerns for Cu interconnects in advanced semiconductor chips. It is caused by Cu atomic diffusion along the electron flow direction due to the momentum exchange with the conducting electrons. Extensive improvements in integration processes and materials selection have been made industry wide to meet the ever challenging requirements for EM resistance [1-3]. However, the aggressive shrinking of Cu lines and vias, compounded with the drive toward lower K dielectric materials for newer technologies has exacerbated EM challenges. The smaller via allows a smaller minimum Cu void size to cause a circuit to fail. Coupling this smaller minimum void size to the relatively wide process variation can cause broad EM failure time distributions, in particular for connections without sufficient liner/via redundancy [4]. Such broadly distributed EM stress failure times make the traditional EM lifetime projection model very pessimistic.

A sound model for failure time distribution is a pre-requisite for accurate projection of EM reliability from accelerated stress data. Not only must the model be able to mathematically fit the experimental data well, but it must also generate the characteristic parameters to properly reflect the physics of the EM process. Traditionally, a 2-parameter lognormal distribution has been used to fit EM failure times. But this fit encounters a difficulty with EM stresses on interconnects without sufficient liner redundancy and via redundancy. This paper discusses how to overcome these difficulties with a more appropriate 3-parameter lognormal distribution fitting.

2- VS 3-PARAMETER LOGNORMAL DISTRIBUTIONS

The 2-parameter lognormal distribution has been most commonly used to characterize EM failures for interconnects including Al, AlCu and Cu in ULSI circuits for the last four decades. It is described as

$$f[t; t_{50}, \sigma] = \frac{1}{t\sqrt{2\pi\sigma^2}} e^{-\left(\mathrm{Log}\left[\frac{t}{t_{50}}\right]\right)^2/2\sigma^2} \tag{1}$$

$$F[t; t_{50}, \sigma] = \int_0^t f[x; t_{50}, \sigma]\, dx \tag{2}$$

$$z = (\mathrm{Log}[t] - \mathrm{Log}[t_{50}])/\sigma \tag{3}$$

where f is the probability density function (PDF) and F is the cumulative distribution function (CDF) of lognormal distribution; t_{50} (the median time to failure) and σ (the standard deviation of the logarithm of failure times) are the characteristic parameters of the distribution (i. e. the two parameters we are referring to), and z is the score function that transforms failure times to the standard Gaussian scale. On the z vs $Log[t]$ probability plot, the ideal single modal distribution is a simple straight line. For some via/line EM structures, failure time distributions have been found to severely deviate from ideal linear behavior. Introduction of two bimodal lognormal distributions was needed to obtain a satisfactory fit [5]. These models were

1) the superposition model, where the overall distribution is the sum of two 2-parameter lognormal distributions,

$$F[t] = p1 * F1[t; t_{501}, \sigma1] + (1 - p1) * F2[t; t_{502}, \sigma2] \tag{4}$$

and 2) the weakest link (or competing risks) model ,

$$F[t] = F1[t; t_{501}, \sigma1] + F2[t; t_{502}, \sigma2] - F1[t; t_{501}, \sigma1] * F2[t; t_{502}, \sigma2] \tag{5}$$

where F is the overall CDF, $F1$ and $F2$ are the CDF for early mode and late mode, respectively, and $p1$ is the fraction of the early mode.

These 2-parameter and bimodal lognormal distribution models have worked seemingly well for aluminum based interconnects. For the advanced Cu interconnects, while they seem still working well for many structural configurations, fits to the data are less satisfactory for structures without sufficient liner and via redundancy.

Conceptually, the 2-parameter lognormal distribution should not be an appropriate model for EM failure times, since EM is a wear-out failure mechanism. As equation (3) indicates, as z becomes smaller and smaller, t approaches zero. This implies that nearly instant EM failure is possible, though the probability may be very low. Although instantaneous failure is possible for defect-induced fails, it seems intuitively evident that it should never happen for wear-out failure mechanisms, such as EM. Since EM failure is caused by metal voiding resulting from atomic diffusion, a critical minimum void size (V_{crit}) is needed to cause an observable resistance increase for the given interconnect structure. This critical void requires a finite time for the metal atoms to diffuse. For instance, the weakest link for damascene Cu interconnect EM is the via/line contact region, where Cu voiding either under or within the via bottom can cause failure. To cause any “significant” resistance increase, the minimum void size (V_{crit}) needs to be large enough to cover most of the via bottom. The minimum time (t_{min}) needed to form this void through Cu atomic diffusion is described by:

$$t_{min} = \frac{V_{crit}}{s v_d} = \frac{V_{crit}}{s\left(\frac{D_{eff}}{kT} Z^* e\rho j\right)} \tag{6}$$

where s is the effective cross section area for Cu atomic flux, v_d is Cu drift velocity, D_{eff} is the effective Cu diffusivity, Z^* is the effective charge number, e is the electronic charge, ρ is the Cu electrical resistivity, j is the stress current density, and kT is the thermal energy. Since V_{crit} is a definite positive volume (it varies with via/line layout under the same stress conditions), t_{min} must be a finite positive number as well. More specifically, there should be a threshold time ($t_{min} > 0$) before which no EM failures can occur. This threshold time is the minimum failure time for a given interconnect under specific stress conditions. The appropriate lognormal distribution for EM failure times with a minimum threshold should be a 3-parameter distribution, which is given by:

$$f[t;\, t_{50},\, \sigma,\, X_0] = \frac{1}{(t-X_0)\sqrt{2\pi\sigma^2}}\, e^{-\left(\mathrm{Log}\left[\frac{t-X_0}{t_{50}-X_0}\right]\right)^2 / 2\sigma^2} \tag{7}$$

$$z = (\mathrm{Log}[t-X_0] - \mathrm{Log}[t_{50}-X_0])\,/\,\sigma \tag{8}$$

where X_0, the third parameter, is the minimum threshold failure time (t_{min}). To emphasize that the minimum threshold failure time is not the failure time observed from the first fail among the samples in a set of stress, but rather is a theoretical limit which the distribution approaches, symbol X_0, rather than t_{min}, is used from now on to represent this minimum threshold failure time. In theory, the earliest failure time approaches this minimum threshold failure time as sample size is increased, but it will never quite reach it for finite sample sizes.

Filippi et al [6] applied equation (8) on EM failure data analysis of Ti-AlCu-Ti interconnect. They pointed out that use of the 2-parameter lognormal distribution leads to the paradoxical results that lower stress-current density or selection of a higher resistance shift criterion for failure could lead to shorter projected lifetimes. They successfully resolved this paradox by using the 3-parameter lognormal distribution.

The basic difference between 2-parameter and 3-parameter lognormal distribution is whether the z vs *Log[t]* line on the probability plot is straight linear or curved downward (for positive X_0). As will be shown in the following discussions, this difference may be hard to distinguish or can be easily noticeable from the experimental data, depending on the parameters of the distributions.

Figure 1 compares the basic functions of the 2-parameter and 3-parameter lognormal distributions. The shift to the right for the 3-parameter lognormal PDF curve (Figure 1(a)) highlights the impact of X_0 (minimum failure time for EM). Figure 1(c) shows the distinguishing feature of z vs *Log[t]* of the 3-parameter lognormal distribution, the bending downwards (for positive X_0) away from the linear as for 2-parameter distribution. Figure 2 presents examples of the impact of X_0 and σ on the differences between the 2-parameter and 3-parameter lognormal distributions. These figures clearly show that both the shape parameter, σ, and the differences between t_{50} and X_0 have significant impact on the difference between the two distributions, especially for the early portion of the distribution. For a CDF as low as 10^{-5} with a tight distribution (σ=0.2, as shown in Figure 2(a)), the difference between 2-parameter (linear) and 3-parameter (curved) lognormal distributions is hard to distinguish. For the same X_0 and t_{50}, the difference between the two distributions become evident for the broader distributions (σ=1.0) for CDF as high as 10^{-3} as shown in Figure 2(b). The failure time distributions for the traditional EM stress data from Al based interconnects are usually tight (σ in the range of 0.1 - 0.3), they seem to follow the 2-parameter lognormal distribution very well for sample sizes in dozens (seemingly good linearity on the probability plot, without severe bending downwards). Even for Cu interconnects with decent redundancy, most of the failure time distributions are not terribly broad, and the deviation from linearity feature in probability plots is not strikingly evident with sample sizes up to hundreds. These experiences may have explained why the application of 2-parameter lognormal distribution model in EM data analysis has not been seriously challenged.

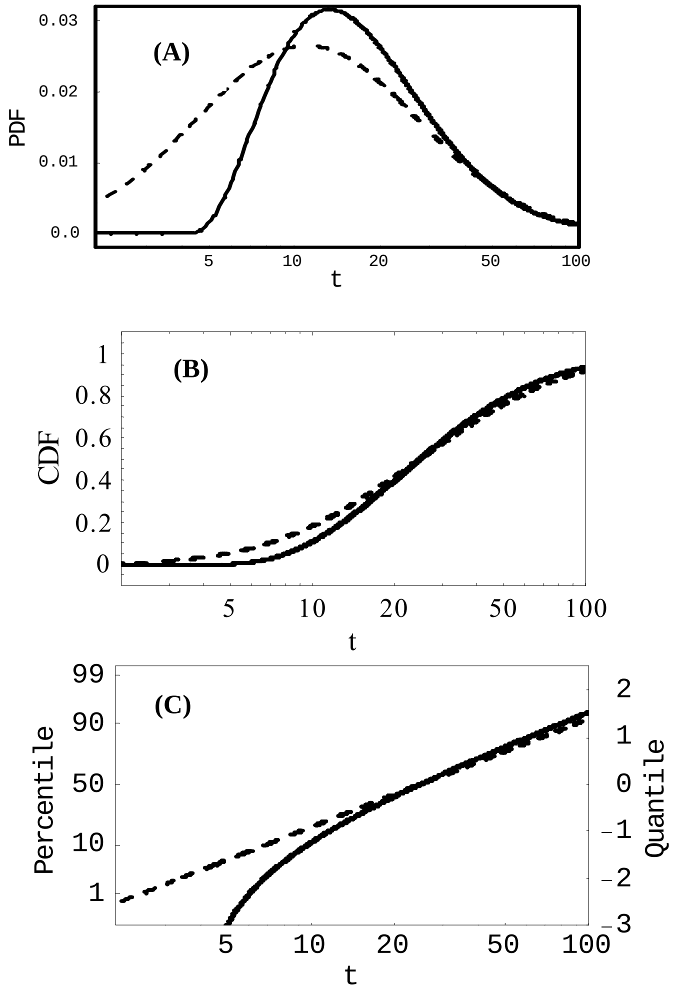

Figure 1. Comparison of 2-parameter and 3-parameter lognormal distributions. The distribution parameters in this example are: t_{50} =25, σ =0.95. The broken lines are for 2-parameter ($X_0 = 0$), and the solid lines are for 3-parameter lognormal ($X_0 = 4$).

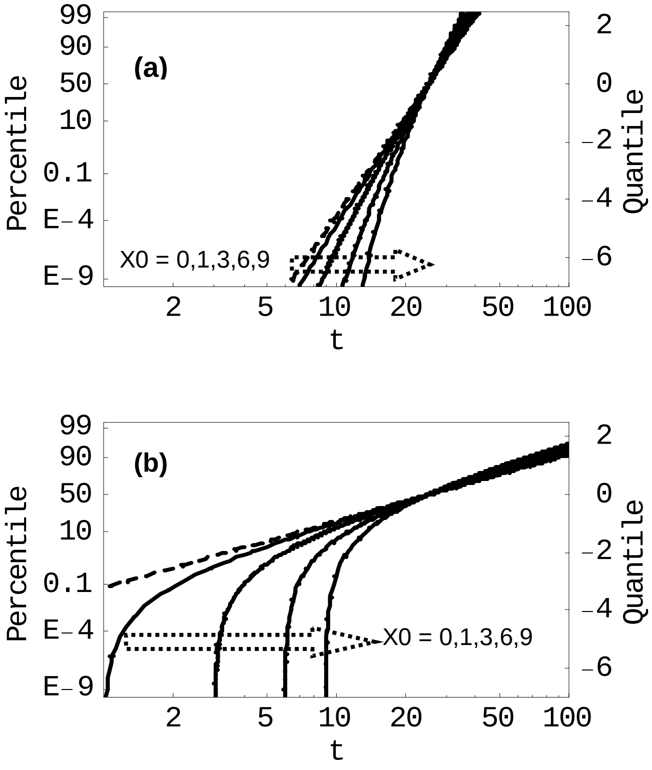

FIGURE 2. IMPACT OF σ AND X_0 ON THE DIFFERENCES BETWEEN 2-PARAMETER AND 3-PARAMETER LOGNORMAL DISTRIBUTION. FIGURE 2(a) ASSUMES $t_{50} = 25$, $\sigma = 0.2$; FIGURE 2(b) ASSUMES $t_{50} = 25$, $\sigma = 1.0$.

VOID SIZE LIMITED EM FAILURE DISTRIBUTIONS IN CU INTERCONNECTS

Failure Time Distribution and Redundancy

As discussed in references [4,7], redundancy (liner redundancy and via redundancy) features in Cu interconnects have substantial effects on EM failure characteristics and distributions. Liner redundancy refers to current carrying-capability of liners when Cu is locally removed from the area, this includes the liners covering the bottom and side walls of vias and lines, and the contact between the via bottom and the liner of the line below (sidewall and/or line end). Via redundancy refers to the number of vias in contact with the lines below or above. The redundancy features determine the critical void size necessary to cause early EM fails (or minimum failure time). For advanced technologies, aggressive shrinking in line and via dimensions reduces the critical void size for failure, and enhances the sensitivity to void shape and location [4]. As a result, the EM failure times become shorter, and the distributions become broader for the cases with less redundancy (both via and liner). Figure 3 shows a few examples of line depletion EM failure time distributions with different redundancies. The stress structure is shown in Figure 4(a) and the via-line contact features at the cathode end are shown in Figure 4(b). The stresses were conducted at 300 ^{o}C at a current density of 2.5MA/cm^{2} in the line, with electrons flowing from V2 down to a 200μm long M2 line. The nominal V2 size is (0.1 x 0.1)μm^{2}, the M2 line width ranges from 0.1μm to 0.7μm. It is clear that with the increase of redundancy, the failure time distributions become tighter and tighter. Cases C and E have both good via redundancy (the maximum number of vias along line width in one row) and liner redundancy (all vias contact the line end liner and the outer vias also contact the M2 sidewall liners), and both show fairly tight distributions and longer times for early fails. Though case A in Figure 4 only has one via, there is good contact between the via bottom and the M2 sidewall liners, because the M2 line has the same width as the V2 via. This good liner redundancy requires a deeper void to cause failure (larger relative critical void size and longer minimum failure time), and to exhibit a steady, gradual resistance increase after the initial resistance jump. Consequently, it has the tightest failure time distribution among the 5 examples in Figure 3. Furthermore, since the stress current was chosen based on the current density in the M2 line, the current crowding at the V2/M2 interface for Case A is also the lowest. Case D has less than the maximum number of vias allowed along the line width and the outer vias do not contact the M2 sidewall liners below. It has less relative via and liner redundancy than Cases A, C, and E, so its failure distribution is broader. Case B has very poor via and liner redundancy (single via on a wider line). As a result, it has the broadest failure distribution and shortest early failure times.

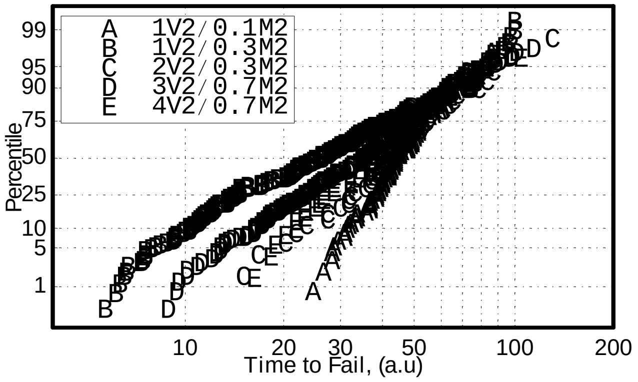

FIGURE 3. EXAMPLES OF LINE DEPLETION EM FAILURE DISTRIBUTIONS – THE DISTRIBUTION BECOMES BROADER (HIGHER σ) WITH THE REDUCTION OF THE REDUNDANCY (BOTH LINER AND VIA).

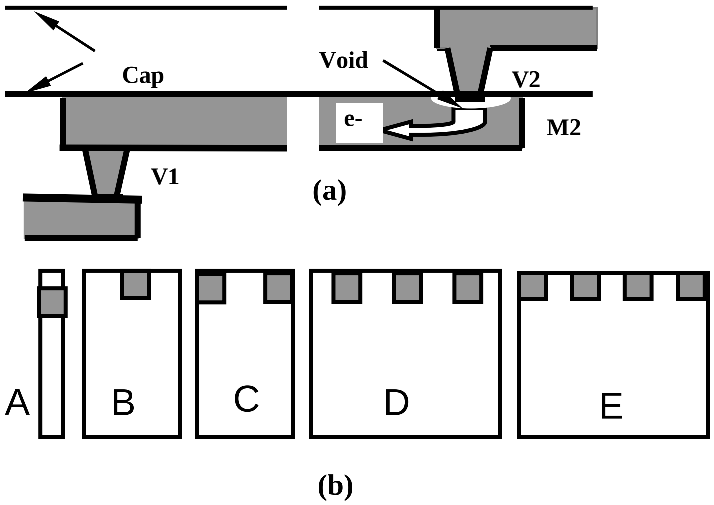

FIGURE 4. SCHEMATIC OF THE STRUCTURES FOR EM STRESSES SHOWN IN FIGURE 3. (a) SCHEMATIC OF THE CROSS SECTION OF THE STRUCTURE; (b) VIA/LINE CONTACT OF THE CATHODE END.

A legitimate question here is whether the local current crowding has an impact when comparing the failure time distributions above. It is true that the severity of current crowding at the via/line contact interface is different for the cases discussed above. Experiments were conducted on case-B at different stress currents. Lower stress current causes failure times to lengthen, but does not alter the distribution shape significantly. This suggests that the local current crowding should not be a major concern when comparing the above failure distributions under current stress conditions.

Issues with 2-Parameter Lognormal Distribution

Figure 5 shows the fitting lines of the data from Figure 3 with 2-parameter lognormal distributions. As the redundancy (via and via/line liner contact) decreases, the failure distribution becomes wider and the quality of the fitting becomes poorer. For structures having good redundancy, the tight sigma allows good fitting of the data with single 2-parameter lognormal distributions (A, C and E). However, for Cases B and D, a single 2-parameter lognormal distribution does not fit the data well; *z* is clearly not a linear function of *Log[t]* on the probability plot. It is clear that the 2-parameter lognormal distribution cannot accommodate the bending down of the early portion of these broader distributions. It severely overestimates the EM susceptibility. Recalling the comparisons made in Figures 1(c), 2(a) and 2(b), these experimental data suggest that the 3-parameter lognormal distribution is more appropriate for fitting failure time distributions of such cases.

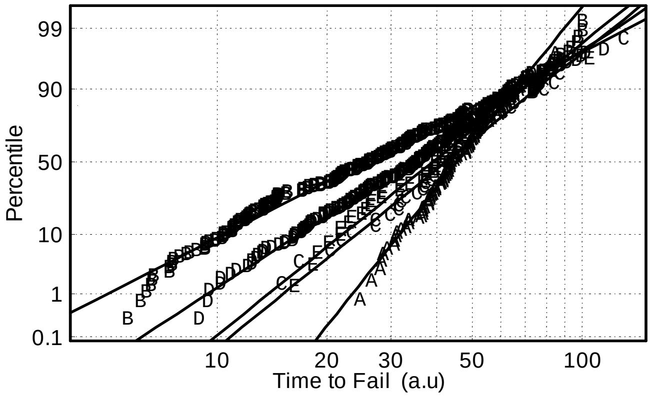

FIGURE 5. FITS OF THE DATA FROM FIGURE 3 WITH 2-PARAMETER LOGNORMAL DISTRIBUTION

Application of 3-Parameter Lognormal Distribution

Figure 6 shows the fitting curves of the experimental data with 3-parameter lognormal distribution described by equations (7) and (8). Visual inspection clearly shows that the 3-parameter lognormal distribution produces a much better fit than the 2-parameter lognormal, especially for the early portion of the broader distributions (B and D). Quantitative analysis of goodness of fit for these data will be published separately [8]. For EM life time projection, the early portion of the distribution is the most important part, and those early fails are the real reliability concerns.

To further validate application the 3-parm lognormal to these EM failure distributions, physical failure analyses were performed on samples from stresses of case B with various failure times as shown in Figure 7(a). Figure 7(b) presents the cross section pictures of these samples. They confirmed that all failures follow the same mechanism with void under the via. The only difference between the early and late fails is the void size and shape. The very early fail (1) has a very thin slit-like void right underneath the via. The late fail (4) has a larger void, and it seems that the void did not start right underneath the via, but adjacent to the via and eventually grew to the via bottom to cause the fail. These SEM pictures demonstrate that the failure times are determined by void sizes and shapes. The root cause of the wide distribution is the sensitivity of the failure times to the size and shape of the void needed to cause failure, due to the lack of (or insufficient) redundancy. With better redundancy, the minimum void size to cause failure is larger, the failure time is less sensitive to void shape and location [4], and consequently, tighter failure time distributions will be achieved.

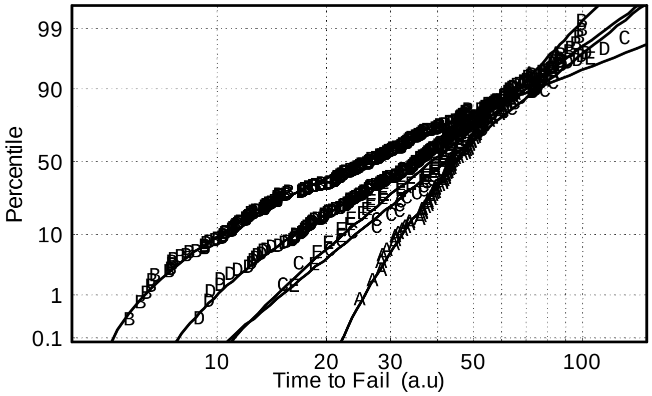

FIGURE 6. FITS OF THE RESULTS FROM FIGURE 3 WITH 3-PARAMETER LOGNORMAL DISTRIBUTION

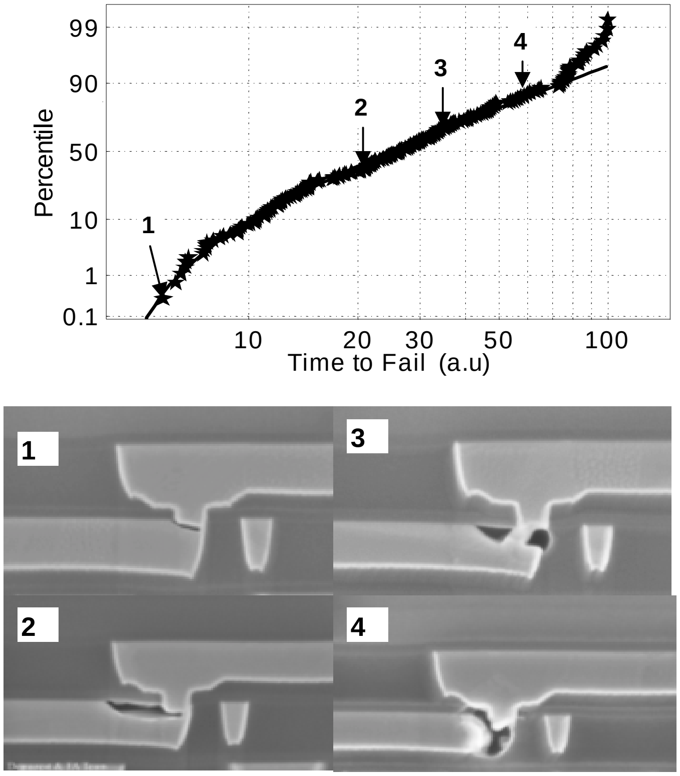

FIGURE 7. EM FAILURE CHARACTERISTICS OF CASE B. (A) (TOP) – THE FAILURE TIME DISTRIBUTION AND PFA SAMPLE LOCATIONS; (B) (BOTTOM) – CROSS SECTIONS OF THE REPRESENTATIVE SAMPLES SHOWING THE VOID FEATURES.

For any distribution to be used in reliability modeling (such as EM), mathematically fitting the experimental data well is only one of the necessary requirements. As we stated earlier, the parameter values derived from the model have to make physical sense for the failure mechanism under study. Since the fitting parameters will be used for lifetime projection, it is very important to ensure that the parameters obtained are stable. One aspect of the parameter stability is the variation with data censoring. Analysis of the accelerated lifetime stress data often involves right censoring. The appropriate censoring should enable the fitting to include all the data points belonging to the same distribution and to exclude those points not belonging to the distribution. Some degree of variation for the estimated parameter values is expected with different data censoring. However, significant changes in estimated parameter values with minor data censoring variation should raise questions on the validity of the assumed failure model. For the failure time distribution of case B (symbol B in Fig.3), the very late portion of the population shows a different trend from the rest of the population. Failure analysis further confirmed that those are the failures with voids in the line, away from the via, differing from the rest, which all had void under the via. They should be censored prior to fitting. Figure 8 shows the values of the mean of the squared residuals (that serves as a goodness of fit measure) as a function of the right censoring time. The error increases dramatically when failure time data beyond 75 is included in the distribution prior to fitting. Table I lists the fitting parameters obtained at different right censoring times. Table I also compares the estimated parameter values of case B with nonlinear regression and Maximum Likelihood Estimation (MLE) with various right data censoring. Both Figure 8 and Table I suggest that 75 is the appropriate time for the right data censoring of this example. With data censoring from 50 to 75, the estimated parameter values are fairly stable. Detailed discussions on the parameter estimation and confidence bound analyses are beyond the scope of this paper, and they are planned to be published separately [8].

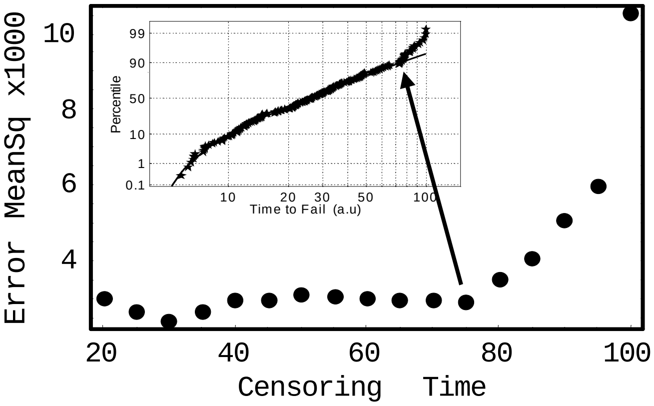

FIGURE 8. FITTING RESIDUAL VARIATIONS WITH RIGHT CENSORING TIME FOR CASE B. NOTE THAT THE ERROR INCREASES RAPIDLY IF FAILS BEYOND 75 ARE INCLUDED IN THE DISTRIBUTION PRIOR TO FITTING.

It should be kept in mind that the EM failure time distributions discussed above and in what follows are determined by the interconnect redundancy features; they are not in any way related to gross process defects and/or gross non-uniformities across a wafer.

Via Size Impact on Minimum Failure Time

While it is obvious that the 3-parameter lognormal distribution can fit the failure time data better for line depletion EM with insufficient redundancy, additional validation is provided by scaling of the estimated parameter values. Many factors can affect minimum failure times, such as void growth rate (by stress conditions and process/integration conditions), layout characteristics (redundancy and passive Cu reservoir) and geometric dimensions. As has been discussed earlier (see equation (6)), the minimum void size to cause an early failure should be proportional to the length of the via/line contact area, or the via length for structures with the same via/line width ratio. At the same stress conditions and for the same integration scheme, a larger via should result in a longer minimum failure time (X_0), since more time is needed to form a thin slit void that covers the entire via bottom. Figure 9 shows failure time distributions for three technology nodes with different via sizes for cases of line depletion EM with similar redundancy features (single via contacts a wide line below). Table II lists the details of the via sizes, integration, estimated EM failure time distribution parameters. Comparing the data with similar integration schemes and via to line width ratio, the minimum failure time (X_0) scales fairly well with the via size.

Figure 9 also shows the comparison between 3-parameter and 2-parameter fits; the details are listed in Table II. The 3-parameter lognormal distribution clearly fits the data better for all these cases.

Table I. Estimated Parameters from Failure Time Distribution of Case B

Right Censoring Time	X_0 (a.u)		t_{50} (a.u)		σ	
	MLE	Regress	MLE	Regress	MLE	Regress
100	2.94	2.37	24.94	25.16	0.82	0.81
90	3.36	3.17	24.85	25.08	0.85	0.87
80	3.64	3.58	24.84	25.05	0.88	0.90
75	4.16	3.74	24.90	25.10	0.94	0.92
70	4.04	3.77	24.86	25.10	0.93	0.93
60	4.09	3.81	24.87	25.11	0.93	0.93
50	3.87	3.78	24.76	25.16	0.90	0.93
40	4.29	4.10	24.98	25.46	0.96	0.98
30	4.62	4.56	25.49	26.27	1.03	1.07

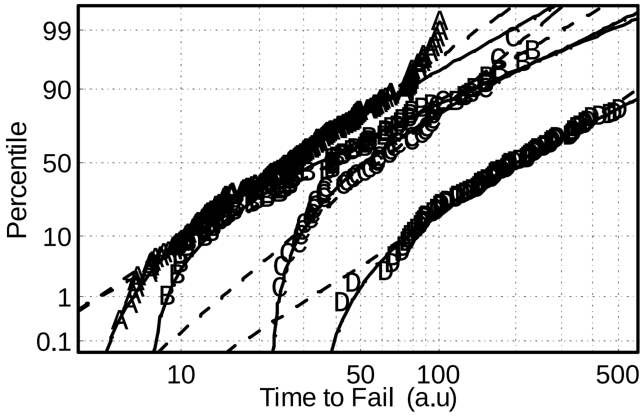

FIGURE 9. EXAMPLES OF EM FAIL TIME DISTRIBUTIONS FOR STRUCTURES WITH SIMILAR POOR REDUNDANCY FROM 3 TECHNOLOGY NODES. SOLID LINES ARE FOR 3-PARAMETER LOGNORMAL FITTING, BROKEN LINES ARE FOR 2-PARAMETER LOGNORMAL FITTING. A – 65nm NODE; B – 65nm NODE FAT WIRE; C – 90nm NODE; D – 180nm NODE. SEE TABLE II FOR DETAILS.

LIFE TIME PROJECTION WITH 3-PARAMETER LOGNORMAL DISTRIBUTION

To prevent EM induced interconnect failures, circuit designers are usually given guidelines on the maximum current density allowed in the metal lines and vias based on the product application specifications. JEDEC [9] recommends general procedures to project the allowed design current from accelerated stress data. The two key parameters from the stress data used for the projection are t_{50} (the median failure time) and σ (the distribution shape factor). For general semiconductor chip applications, the allowed accumulated end life failure probability (value of the CDF) per connection is usually very low (typically less than 10^{-11}). Projection from the stress sample size (usually in the order of hundred or less, where the minimum CDF value is about 10^{-2}) to the general population scale (10^{-11}) is about 9 orders of magnitude, and the projected lifetime becomes extremely sensitive to the variation of σ, especially for distributions with high σ values. As discussed in previous sections, for interconnect structures with poor redundancy, this approach can result in very pessimistic lifetime projections. When a 3-parameter lognormal distribution is applied, the dependency on σ for the lifetime projection can be significantly alleviated.

Table II. Minimum Failure Time Variation with Via Size Scaling

Technology Node	65nm (thin wire)	65nm (fat wire)	90nm	180nm
ILD	SiCOH	SiCOH	FSG	SiO_2
Via size μm^2	0.1x0.1	0.2x0.2	0.14x0.14	0.28x0.28
Line width	0.3µm	0.6µm	0.28µm	0.76µm
3-parm:X_0	4.16	7.3	21.9	33.5
3-parm: t_{50}	24.9	33.8	51.8	200.7
3-parm:σ	0.94	1.23	1.02	1.08
Goodness of fit: 3-parm	0.003	0.013	0.017	0.006
Goodness of fit: 2-parm	0.009	0.018	0.036	0.015

Note: 1) The Cu/cap interface for 65nm differs from that for the earlier technologies (90nm and 180nm). The Cap/Cu interface of 90nm samples stressed is similar to that of 180nm.
2) The goodness of fit is defined as the mean of the squared residuals.

Based on Black's acceleration model for electromigration [10]:

$$MTTF = Aj^{-n} e^{\frac{\Delta H}{kT}}. \tag{9}$$

In equation (9), *MTTF* is the median time to fail, *A* and *n* are constants, and *j* is the stress current density, ΔH is the activation energy for metal diffusion, *k* is Boltzmann constant, *T* is the interconnect temperature (in K).

We have shown in previous sections that the physical failure mechanism is the same for the fails across the entire uniform distribution range. This suggests that the relation (9) is satisfied not only for the median of the failure time, but also for an arbitrary quantile of the failure time distribution, where A is determined by the order of the quantile. In light of this fact, the maximum allowable use current density (J_{use}) for a given interconnect may be calculated as:

$$J_{use} = J_s \left(\frac{t_f}{t_{EOL}}\right)^{(1/n)} \mathrm{Exp}\left[\frac{\Delta H}{nk}\left(\frac{1}{T} - \frac{1}{T_s}\right)\right] \tag{10}$$

where J_s is the stress current density, T_s is the stress temperature (in K). *n, k,* ΔH and *T* have the same meaning as in equation (9), and t_{EOL} is the designed quantile of a given order (e.g., median) of the product lifetime. t_f is the projected quantile of the same order of the failure time distribution under stress conditions; for the lognormal distribution it is given by:

$$t_f = \mathrm{Exp}[U_1 \sigma + \mathrm{Ln}(t_{50} - X_0)] + X_0 \tag{11}$$

where U_1 is the quantile of interest for the given CDF at product's end of life, expressed on the standard Gaussian scale. For the distributions with high σ, the term $+X_0$ in the t_f calculation (equation (11)) can play a crucial role for the lifetime projection, and can give a major relief compared to 2-parameter lognormal distribution model. It is also this term that reduces the lifetime projection dependency on the shape factor, σ. Furthermore, if the quantile of interest t_{EOL} is the left endpoint of the failure time distribution (i.e., quantile of order zero), then U_1 approaches negative infinity, resulting in

$$t_f = X_0 \tag{12}$$

and the J_{use} projection is simplified as

$$J_{use} = J_s \left(\frac{X_0}{t_{EOL}}\right)^{1/n} \exp\left[\frac{\Delta H}{nk}\left(\frac{1}{T} - \frac{1}{T_s}\right)\right] \tag{13}$$

which makes the J_{use} projection completely independent of t_{50} and σ. The sole parameter needed from the stress data is the minimum failure time, X_0, in addition to the kinetic parameters ΔH and *n*.

Equations (10) and (13) contain the same kinetic parameters, *n* and ΔH, as for the 2-parameter lognormal model. To produce an accurate lifetime or maximum allowed current density projection, these kinetic parameters should be generated by the 3-parameter lognormal modeling.

Special attention needs to be paid to joule heating issues when applying equation (10) or equation (13) for J_{use} projection. In both equations, *T* is the temperature of the interconnect under consideration. It may differ from the nominal chip operating temperature (T_u), depending on the level of joule heating of the interconnect itself, by its neighbors or both. When joule heating is not negligible, $T = T_u + \Delta T$. ΔT is the temperature rise of the interconnect caused by joule heating. The details on how to include joule heating in the lifetime projection are discussed by Hunter [11,12] and Li, et al [13].

DISCUSSIONS

Sample Size Considerations

As shown in Figures 1(c), 2(a) and 2(b) and already alluded to above, the graphical difference between 2-parameter and 3-parameter lognormal distribution on the probability plot is that the curve for 3-parameter fit bends down (for positive X_0) at the left (early failure), while the 2-parameter curve is a straight line. For tight distributions (low σ), this downward bending may not become apparent on the probability plot, except for very large sample sizes. Applying 2-parameter lognormal distribution for EM lifetime projection to these cases may not be prohibitively conservative, and usually projects reasonably accurate values at use conditions. However, for stress data with very broad distributions (high σ), like case B in Figure 3, the inherent limitations of using a 2-parameter lognormal distribution in EM data analysis become apparent. The projected lifetime becomes prohibitively conservative, and does not appropriately reflect the technology capability. While 3-parameter lognormal distribution makes sense both mathematically and physically for EM data analysis, it usually requires a much larger sample size for reasonable parameter estimation. Figure 10 gives one example of variation of estimated parameter values from Monte Carlo simulations. The input parameter values for the 3-parameter lognormal function are: $X_0 = 4.16$, $t_{50} = 24.9$, $\sigma = 0.94$. 200 simulations were conducted for each sample size. For this specific example, the sample size needs to be 240 or greater to ensure 90%

probability that the estimated X_0 falls within the range of 3.1 – 5.4 (the input value is 4.16).

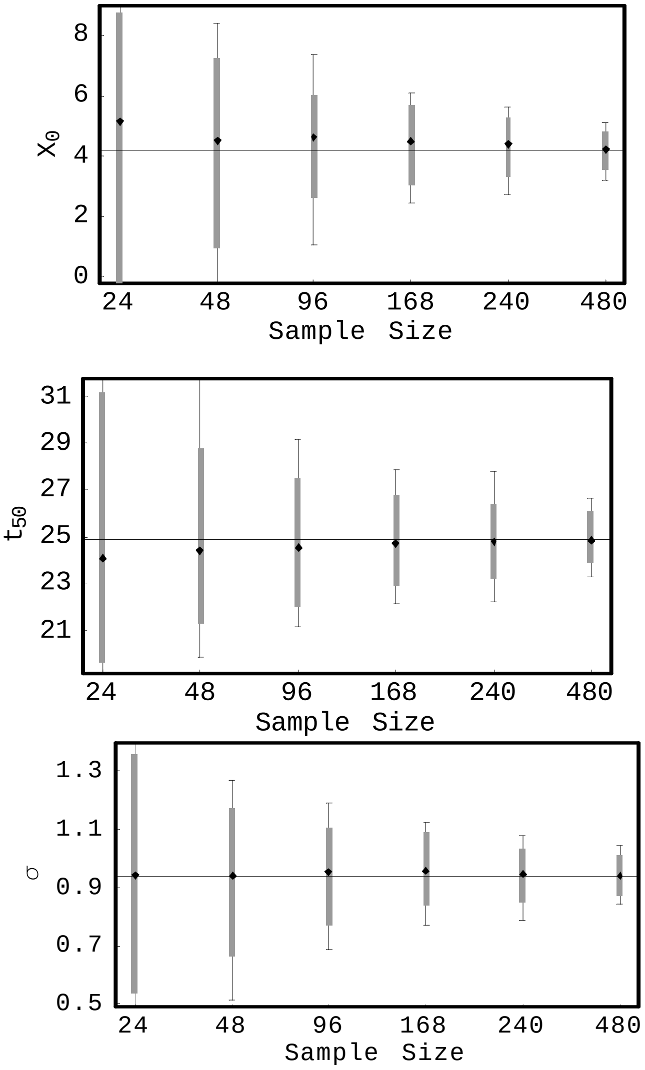

Figure 10. Monte Carlo simulations showing estimated parameter variation with sample size. The gray bars cover 10 – 90 percentiles, and the vertical line bars cover 2.5 – 97.5 percentiles. The black dot is for the 50 percentile. The horizontal line is the input value (the expected) for the simulation. The 3-parameter lognormal function used is: $X_0 = 4.16$, $t_{50} = 24.9$, $\sigma = 0.94$. For each sample size, 200 simulation runs were conducted.

Redundancy Impact on Minimum Failure Times

Equations (11) and (12) show the importance of X_0 values on EM lifetime projection. It seems straightforward from equation (6) to model X_0, if the Cu diffusion parameters are known. In reality, V_{crit} is a very sensitive function of interconnect redundancy (both liner redundancy and via redundancy). Similar to the discussions made in reference [4], when the vias have good contact with the line liner below for line depletion EM, the void underneath the via needs to grow deeper into the Cu line to cause a resistance increase greater than a pre-set EM failure criterion. Furthermore, the details of via/line liner contact (such as contacting line end liner, sidewall liner, and whether or not there is a line extension under the via), and the liner thickness and electrical properties all affect V_{crit}. For via redundancy, it is easy to understand that with more vias contacting the line, greater V_{crit} is needed to observe the resistance increase needed to cause an EM failure. The layout arrangement of the redundant vias (number of rows, columns, and their spaces, with and without line extensions, etc) will also have significant impact on V_{crit}. Therefore, these variables all should be taken into consideration for maximum allowed design current guidelines.

SUMMARY

Using an appropriate distribution model is very important for accurate reliability projections. Considering the wear-out nature of the EM failure mechanism, the 3-parameter lognormal should be more appropriate than the commonly used 2-parameter lognormal for EM failure time distributions. For the EM stress data with reasonably tight distributions, the 2-parameter lognormal distribution can still produce acceptable lifetime projections for practical purposes. However, for EM stress data with broad distributions (high σ), the 2-parameter lognormal model may become severely conservative and fails to reflect the behavior of early failure. The 3-parameter lognormal distribution can better fit the stress data and produce more accurate projections for the interconnect EM resistance provided an adequate sample size is used.

ACKNOWLEDGEMENTS

The authors acknowledge the management support from Dinesh Badami; stress support from Bill Hastings, Chad Burke and Randy Austin; and PFA support from James Demarest and his team. Valuable discussions with the IBM BEOL reliability team are also highly appreciated.

REFERENCES

[1]. C. Bruynseraede, et, al; “The impact of scaling on interconnect reliability”, Proc. 43rd Annual Reliability Phys Symp. (IRPS), IEEE, 2005, p7-17.

[2]. B. Li, et al; “Reliability challenges for Cu interconnects”, Microelectronics Reliability, 44(3), 2004, p365-380

[3]. C.-K. Hu, et al; “Effect of overlayers on electromigration reliability improvement for Cu/low K interconnects”, Proc. 42rd Annual Reliability Phys Symp. (IRPS), IEEE, 2004, p222-227.

[4]. B. Li, et al, “Impact of via-line contact on Cu interconnect EM performance”, Proc. 43rd Annual Reliability Phys Symp. (IRPS), IEEE, 2005, p24-30.

[5]. A. H. Fischer, et al; “Experimental data and statistical models for bimodal EM failures”, Proc. 38th Annual Reliability Phys Symp. (IRPS), IEEE, 2000, p359 -363.

[6]. R. G. Filippi, et al; “Paradoxical predictions and minimum failure time in electromigration”, Appl. Phys. Lett., 66(16), 1995,pP1897-1899

[7]. B. Li, et al, “Line depletion EM characteristics of Cu interconnects”, IEEE Trans. On Devices & Materials Reliability, 4(1), 2004, p80-85.

[8] B. Li, et al, “Application of 3-parameter lognormal distribution in EM data analysis,” to be published.

[9]. JEDEC/FAS Joint Publication, JP001, Foundry process qualification guidelines, Setp. 2002.

[10]. J. R. Black, “Mass Transport of Aluminum by momentum exchange with conducting electrons”, Proc. 43rd Annual Reliability Phys Symp. (IRPS), IEEE, 2005, p1-6 (Reprinted from IRPS 1968).

[11]. W. Hunter, Self-consistent solutions for allowed interconnect current density – Part I: Implication for technology evolution", IEEE Trans. Electronic Devices, 44(2), 1997, p304-309.
[12]. W. Hunter, Self-consistent solutions for allowed interconnect current density – Part II: Application to design guidelines", IEEE Trans. Electronic Devices, 44(2), 1997, p310-316.
[13]. B. Li, et al; "Thermal and EM challenges for advanced interconnects", 2004 IEEE International Integrated Reliability Workshop Final Report, p46-51.