**RC23680 (W0507-213) July 29, 2005**
**Mathematics**

# IBM Research Report

## Application of Three-Parameter Lognormal Distribution in EM Data Analysis

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

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

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

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

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

# Application of Three-parameter Lognormal Distribution in EM Data Analysis

**Baozhen Li[1)], Emmanuel Yashchin[3)], Cathryn Christiansen[1)], Jason Gill[1)], Ronald Filippi[2)] and Timothy Sullivan[1)]**
**IBM Systems and Technology Group**
**1) Essex Junction, VT 05452; 2) Hopewell Junction, NY 12533**
**3) IBM Research Division, Yorktown Heights, NY 10598**
Lib@us.ibm.com

## Abstract

Three-parameter lognormal distribution has been demonstrated for applications in electromigration data analysis, especially for Cu interconnect structures with insufficient redundancy. Examples are given on estimating parameter values from experimental data using the maximum likelihood method. Detailed analyses are presented on confidence bound estimations of the parameters and their propagation to lifetime projections.

KEY WORDS: Electromigration, Cu interconnects, voids, reliability, failure distribution, lognormal, maximum likelihood, confidence bound.

## 1. Introduction

Electromigration (EM) is a wear out failure mechanism for on-chip metal interconnects in ULSI (Ultra Large Scale Integrate) circuits. The root cause for EM failures is the interconnect voiding from metal atomic diffusion driven by the "electron wind". The weakest link for EM failures in Cu interconnects is the via/line interface (either within or below the via), where a thin slit like-void can cause a catastrophic circuit failure [1-4]. For advanced semiconductor technologies, the interconnect features, especially the Cu damascene vias, are becoming smaller and smaller, down to a fraction of a tenth of a micrometer. This via size shrinking results in a smaller critical void size, and consequently a shorter time, to cause a substantial resistance increase or open circuit failures. Furthermore, for interconnect structures with poor redundancy (both via and liner redundancy), EM failure times are very sensitive to void location and shape, and show broader distributions [1,5]. Traditional EM analysis with 2-parameter lognormal fitting leads to unacceptably short lifetime projections for such failure distributions. As discussed by the current authors [5], a more appropriate three-parameter lognormal distribution model can better fit the data and produce more reasonable projections. In this paper, we present detailed discussions on the application of a three-parameter lognormal model in EM data analysis, including estimating parameter values and their confidence bounds from experimental data using maximum likelihood and related methods. Propagation of the errors from these estimated parameters to the projected lifetime will also be discussed.

## 2. Three-parameter Lognormal Distribution in EM Modeling

Traditionally, EM failure times have been analyzed by lognormal distributions with two characteristic parameters, $t_{50}$ (the median time to failure) and σ (the shape factor, or the standard deviation of the logarithm of failure times) [6,7,8]. Since EM is a wear out failure mechanism, a minimum void with finite size is needed to cause an interconnect resistance increase above a preset failure criterion. To form this critical void, certain amount of Cu has to diffuse away from the divergence spot, which requires definite time. Therefore, unlike defect-caused failures, there is no instant EM induced failure; a finite minimum failure time is needed for an EM failure. For distributions with a minimum threshold failure time, $X_0$ , a 3-parameter lognormal fit has been found to be more appropriate [5]. The probability density function (PDF) $f(t)$ of the three-parameter lognormal distribution is given by:

$$f(t;t_{50},\sigma,X_0)=\frac{1}{(t-X_0)\sqrt{2\pi\sigma^2}}\exp\left[-\frac{\ln^2\left(\frac{t-X_0}{t_{50}-X_0}\right)}{2\sigma^2}\right],\quad t>X_0 \qquad (1)$$

Formula (1) uses special parametrization that is convenient for work with reliability data. In what follows we will denote by T the random variable representing the failure time distributed in accordance with (1).

The major visual difference between a two-parameter and a three-parameter lognormal distribution is a straight line vs. a downward bending curve at the early fail portion (for a positive $X_0$) in the normal probability plot for logarithms of failure times [5,9]. The difference between these two distributions depends heavily on the shape factor, σ, and the difference between $t_{50}$ and $X_0$. As shown in Figure 1(a), for a tight distribution with σ = 0.2, the difference is hardly distinguishable between three-parameter and two-parameter lognormal fits for a sample size as large as 1000. When σ is increased to 1.0, however, for the same $t_{50}$ and $X_0$ values, the difference between the two fits becomes very evident for a sample size of only 100; see Figure 1(b).

For Cu interconnects, due to the damascene process, the redundancy feature of the structures has significant effect on EM distributions [1]. Good redundancy (for both liner and via) usually leads to tight EM distributions, while poor redundancy results in broad distributions. Consequently, straight lines are seen on the probability plots for log($T$) for EM data from structures with good redundancy, while a downward curvature in the early fail portion is observed for EM structures with poor redundancy. Figure 2 shows four examples of EM stress data using similar structures with poor redundancy from 3 technology nodes. The common feature for these EM structures is a single via contacting a wider Cu line below with electrons flowing from the via down to the line. The line widths are 2-3 times of the width of the via. Regardless the via size and integration details, all these data show similar distribution shapes, with early portion of the probability plot bending down wards. Table I lists the goodness of fit comparison between two-parameter and three-parameter lognormal distributions. It is clear that only the three-parameter fitting can reflect the characteristics of the early fail population, which is crucial for lifetime projection.

### 3. Parameter Estimations with Maximum Likelihood

Maximum Likelihood Estimation (MLE) is the most widely used technique in statistics, due to its desirable efficiency and the least biased propositions [10]. The likelihood function for a three-parameter lognormal sample with single right data censoring is given by:

$$L(t_{(1)},t_{(2)},...,t_{(m)};t_{50},\sigma,X_0)=\frac{n!}{(n-m)!}\prod_{i=1}^{m}\frac{1}{\sigma\sqrt{2\pi}(t_{(i)}-X_0)}\exp\left[-\sum_{i=1}^{m}\frac{\ln^2\left(\frac{t_{(i)}-X_0}{t_{50}-X_0}\right)}{2\sigma^2}\right]\times$$
$$\times\left\{1-F\left[\frac{1}{\sigma}\ln\left(\frac{t_c-X_0}{t_{50}-X_0}\right)\right]\right\}^{n-m},\quad t_c\geq t_{(m)}\geq t_{(m-1)}\geq...\geq t_{(1)}>X_0, \qquad (2)$$

where $n$ is the total number of samples under stress, $m$ is the number of data points to be included in the fitting, $\{t_{(1)},t_{(2)},...,t_{(m)}\}$ are the ordered failure times. Our censoring here of type 1, i.e., the censoring time $t_c$ is established a-priori; so that the experiment is considered terminated at time $t_c$ and failure times exceeding $t_c$ are either not observed or considered irrelevant. One can also produce a similar expression for type 2 censoring, where the experiment is discontinued after observing the $m$-th failure; in this case $t_c$ is simply replaced by $t_{(m)}$ in (2). $F$ is the cumulative distribution function (CDF) of the three-parameter lognormal distribution corresponding to the density (1). The estimation of parameters ($t_{50}$, $\sigma$, $X_0$) involves maximizing the likelihood function (2) as a function of these three parameters. In practice, it is the logarithm of (2), not the likelihood function (2) itself, that is maximized through solving the following equations simultaneously:

$$\frac{\partial\log(L)}{\partial X_0}=0, \qquad (3)$$
$$\frac{\partial\log(L)}{\partial t_{50}}=0, \qquad (4)$$
$$\frac{\partial\log(L)}{\partial\sigma}=0, \qquad (5)$$

One can show that a unique solution of the equations (3)-(5) exists and it indeed corresponds to the maximum of log-likelihood. This solution defines the maximum likelihood estimates (MLE's); we will denote it by $(\hat{t}_{50}, \hat{\sigma}, \hat{X}_0)$. Discussions of the detailed mathematical procedures on how to solve these equations are beyond the scope of this paper. Thanks to today's computing technology, satisfactory solutions should be fairly easily achievable for EM stress data with reasonable sample size and quality. Examples of three-parametric lognormal estimation based on likelihood maximization can be found in [13].

Data censoring is often necessary for EM data analysis, either because stress intensities have to be kept reasonably low (in order for data to stay within the envelope of the acceleration model) and so the life test is terminated before all samples fail, or because of the presence of multiple failure mechanisms. For accurate parameter estimation, the fitting should include all the data points belonging to the population under study, and exclude the points not belonging to this population. The censoring point may be determined by evaluating the fitting residuals

$$z_i = \frac{1}{\hat{\sigma}} \ln\left(\frac{t_{(i)} - \hat{X}_0}{\hat{t}_{50} - \hat{X}_0}\right), \quad i = 1, 2, ..., m, \tag{6}$$

or the parameter value stability (both should give the same answer). Within a uniform population, the residuals $\{z_i\}$ should correspond (approximately) to a sequence of independent standard normal random variables. Therefore, the mean of the squared fitting residuals should not change substantially with right data censoring variation; but it will increase fairly rapidly when the data points not belonging to the population are included in the fitting. Figure 3 illustrates the goodness of fit (mean of the squared fitting residual) of data A from Figure 2, which corresponds to the experimental data with electrons flowing from a $(0.1 \times 0.1)\mu m^2$ V2 down to a 0.3μm M2 line. The mean of the squared fitting residuals is fairly constant for data censoring from 40 to 75, it starts to increase rapidly after 75 (Figure 3(b)). Visual inspection of the probability plot in Figure 3(a) clearly shows that the population beyond 75 looks different from the rest. This was further confirmed by failure analysis: the samples that failed before 75 had a void directly under the via, while samples that failed later had voids away from the via in the Cu line. For this example, $t_c$=75 is the appropriate censoring time. Figure 3 also shows the variation of the estimated parameter values of $X_0$, $\sigma$ and $t_{50}$. Both $X_0$ and σ track the variations of the goodness of fit well – fairly constant in the censoring range of 40-75, and starting to decrease rapidly beyond 75. The estimated $t_{50}$ does not change significantly as the function of censoring policy. From these results, the mean of the squared fitting residuals proves to be a good indicator for determining appropriate data censoring.

## 4. Estimation of Minimum Failure Time, $X_0$

As will be shown in the later sections, the minimum possible failure time $X_0$ is the most important parameter for EM lifetime projection for structures with poor redundancy. Smaller via size in the newer technologies results in shorter minimum failure times, which makes the issue of estimation of $X_0$ even more important. An MLE point estimate $\hat{X}_0$ of this parameter is obtained by solving the equations (3)-(5). This estimate, however, suffers from a substantial upward bias, which could lead to overly optimistic lifetime projections. It is, therefore, important to include a bias-correcting mechanism as part of the estimation procedure. Such a mechanism will be described in this section.

We will first focus on the procedure for obtaining a lower confidence bound for $X_0$ – this problem is generally more important in practice than the problem of point estimation, as decisions are typically based on confidence statements as opposed to point estimates whose variance may turn out to be too high for practical purposes. As an example, consider the stress data of Case A in Figure 2 (a (0.1 x 0.1) $\mu m^2$ via down to a 0.3μm wide Cu line). In this case the MLE estimate of $X_0$ is $\hat{X}_0 = 4.16$, with right data censoring at 75. Based on this estimate, we would like to obtain a 95% lower confidence bound; this amounts to finding a value of Δ for which we could claim that $X_0 \geq 4.16 - \Delta$ with 95% confidence. In general, the confidence bound is a random variable $\hat{X}_0 - \Delta$ that satisfies the property:

$$P\{X_0 \geq \hat{X}_0 - \Delta\} = 1 - \alpha, \tag{7}$$

where $\alpha$ is some number close to zero (α << 1).

One of the most powerful methods for deriving confidence bounds is based on the so-called profile likelihood (e.g., see [11, 12, 13]). This method is based on the fact that if the true value of the parameter is $X_0$, and the maximum likelihood estimates of the other parameters obtained under the assumption that the threshold is $X_0$ are $\hat{t}_{50}^{(X_0)}$ and $\hat{\sigma}^{(X_0)}$, respectively, then the logarithm of the likelihood ratio

$$dL = 2 \times \left\{ \log[L(t_{(1)}, t_{(2)}, ..., t_{(m)}; \hat{t}_{50}, \hat{\sigma}, \hat{X}_0)] - \log[L(t_{(1)}, t_{(2)}, ..., t_{(m)}; \hat{t}_{50}^{(X_0)}, \hat{\sigma}^{(X_0)}, X_0)] \right\} \tag{8}$$

is distributed (asymptotically, as the sample size increases) as a $\chi^2$ random variable with one degree of freedom. In accordance with this statement, derivation of the lower $((1-\alpha) \times 100\%$ confidence bound for $X_0$ involves finding a value of $\Delta$ for which

$$\begin{aligned} dL = 2 \times \left\{ \log[L(t_{(1)}, t_{(2)}, ..., t_{(m)}; \hat{t}_{50}, \hat{\sigma}, \hat{X}_0)] - \log[L(t_{(1)}, t_{(2)}, ..., t_{(m)}; \hat{t}_{50}^{(\hat{X}_0 - \Delta)}, \hat{\sigma}^{(\hat{X}_0 - \Delta)}, \hat{X}_0 - \Delta)] \right\} \\ = \chi^2[1 - 2\alpha, 1] \end{aligned} \tag{9}$$

where $\hat{X}_0$ is the maximum likelihood estimate of $X_0$; the first term in the braces is the maximal value of the log-likelihood log(L), where the maximum is computed with respect to all three parameters ($X_0$, $t_{50}$, and $\sigma$); and the second term in the braces is the maximal value of the log-likelihood, where the constrained maximum is computed by fixing $X_0$ at the value $X_0 = \hat{X}_0 - \Delta$, and maximizing log(L) with respect to the remaining two parameters ($t_{50}$ and $\sigma$). Clearly, this maximization procedure establishes $t_{50}$ and $\sigma$ as implicit functions of $\hat{X}_0 - \Delta$ - and this fact is reflected in the above notation.

For the purpose of EM reliability projection, only the lower bound of $X_0$ is concerned, for instance, for a 95% confidence bound, $\alpha = 0.05$ and the cut-off value for establishing the confidence bound is $\chi^2[(1-2\times 0.05),1] = \chi^2[0.9,1] = 2.70$.

Figure 4 shows the results of $X_0$ confidence bound estimation from profile likelihood based on equation (9). For this example (Case A of Figure 2), at $dL$ =2.7 (95% confidence), the lower bound of $X_0$ is $\hat{X}_0 - \Delta$ = 2.41, or the confidence of $X_0$ being greater than or equal to 2.41 is 95% (note that the maximum likelihood estimate of $X_0$ is 4.16).

Based on the structure of the likelihood function, it can be expected that the MLE would be biased upwards. This, in turn, leads to the lower confidence bound being biased upward, resulting in overly optimistic lifetime projections. It is, therefore, essential to incorporate bias-correction methodology into the estimation process. This bias correction on the estimated parameter values should become especially important for lifetime projections with tight reliability margins and/or smaller sample sizes.

The bias-correction procedure we propose can be viewed as an application of the so-called "parametric bootstrap" technique, with the objective of providing the correct coverage probability of the confidence bound. We define it as follows:

***Procedure A. Bias-corrected estimates and bounds for $X_0$***

(a) obtain the MLE estimates $(\hat{t}_{50}, \hat{\sigma}, \hat{X}_0)$
(b) assume that $(\hat{t}_{50}, \hat{\sigma}, \hat{X}_0)$ are the true values of the parameters. Generate $B$ samples of $m$ first failures out of the test population of $n$ under the assumption that $t_{50} = \hat{t}_{50}, \sigma = \hat{\sigma}, X_0 = \hat{X}_0$; for every sample compute the MLE estimates. This results in a set of $B$ estimates $\{\hat{t}_{50}{}^{(i)}, \hat{\sigma}^{(i)}, \hat{X}_0{}^{(i)}\}, \quad i = 1, 2, ..., B.$
(c) compute the estimated bias:

$$EstBias\ (\hat{X}_0) = \frac{1}{B}\sum_{i=1}^{B} \hat{X}_0{}^{(i)} - \hat{X}_0. \tag{10}$$

(d) set the modified estimate and lower bound for $X_0$ to

$$\begin{aligned} \tilde{X}_0 &= \hat{X}_0 - 0.5 \times EstBias\ (\hat{X}_0), \\ LowerBound(1-\alpha, X_0) &= \hat{X}_0 - \Delta - 0.5 \times EstBias\ (\hat{X}_0), \end{aligned} \tag{11}$$

where $\Delta$ is the solution of (9).

To illustrate the application of the above procedure, consider again the data from case A of Figure 2. The estimated values of the parameters were $\hat{X}_0 = 4.16$, $\hat{t}_{50} = 24.90$, $\hat{\sigma} = 0.94$. Now select B=1000 (the number of bootstrap replications) and generate 1000 censored datasets (based on $n$=240 EM structures on test in each run) from the three-parameter lognormal distribution with parameters $X_0$ = 4.16, $t_{50}$ = 24.90, $\sigma$= 0.94 and censoring time $t_c = 75$. The lower 95% confidence bound for $X_0$ was 4.16-1.75 = 2.41 (see Figure 4).

A total of 1000 sets of data with 240 points in each set (the similar sample size as Case A of Figure 2) were generated. Confidence bound estimation was then performed on each of these 1000 sets of data following the profile likelihood method as described above.

Figure 5 shows the simulation results. There are 943 out of the 1000 values of $X_0 - \Delta$ smaller than or equal to 4.16, which is 7 fewer than what is expected (950), illustrating under-coverage obtained through the straight profile likelihood methodology . In reviewing the estimated $X_0$ values, some upward bias was observed for $X_0$ estimation with MLE at the sample size of 240. For the 1000 simulated data sets, the mean of the estimated $X_0$ values should be very close to the input value of 4.16. But the actual mean value of estimated $X_0$ is 4.31, which indicates a 0.15 upward bias on $X_0$ estimation:

$$EstBias\,(\hat{X}_0) = \frac{1}{B}\sum_{i=1}^{B} \hat{X}_0^{(i)} - \hat{X}_0 = 0.15. \qquad (12)$$

In accordance with Procedure A, we apply the bias correction to both the estimate and the lower bound. In light of (11), this results in

$$\begin{aligned} \tilde{X}_0 &= \hat{X}_0 - 0.5 \times EstBias\,(\hat{X}_0) = 4.16 - 0.15/2 = 4.08, \\ LowerBound(1-\alpha, X_0) &= \hat{X}_0 - \Delta - 0.5 \times EstBias\,(\hat{X}_0) = 2.41 - 0.15/2 = 2.34. \end{aligned} \qquad (13)$$

Simulation study confirms that the proposed estimation method indeed improves the coverage probability of the profile likelihood confidence bounds. In particular, re-evaluation of the original simulation run used in conjunction with Procedure A results in 952 out of 1000 lower bounds being lower or equal to $X_0$ nominal input value, which confirms the 95% coverage probability in this case.

The bias on the estimated parameter values from MLE varies with sample size and parameter value range. As shown in Figure 6, for the distribution with the parameter values discussed in this section, the upward bias for $X_0$ estimation is inversely proportional to the sample size; this is a typical behavior of bias in well-conditioned statistical models. In particular, when the sample size doubles, the bias value decreases to approximately half. At a sample size of 240, the upward bias for $X_0$ is 0.15/2 (the mean of the 1000 estimated $X_0$ nominal is 4.31), while at a sample size of 480, the upward bias of estimated $X_0$ is decreased to 0.07/2 (the mean of the estimated $X_0$ nominal is 4.23), which is about half of that for the sample size of 240.

## 5. Estimation for Lifetime Projections

For purposes of practical application to EM reliability, the lifetime projection is often focused not on lifetime itself, but rather on the maximum allowed use current density ($J_{use}$) for a ***specified*** product lifetime. When the circuit designers follow this $J_{use}$ guideline in their product design, the product is protected from premature EM wear out failures. The projection of $J_{use}$ from stress data is given by a version of Black's equation that specifies the acceleration mechanism with respect to stress conditions [5,6]

$$J_{use} = J_s \left\{ \frac{\exp[U_1\sigma + \log(t_{50} - X_0)] + X_0}{t_{EOL}} \right\}^{1/N} \times \exp\left[ \frac{\Delta H}{Nk}\left( \frac{1}{T} - \frac{1}{T_s} \right) \right], \qquad (14)$$

where $\Delta H$ is the activation energy for Cu diffusion, $N$ is a material-specific constant, $k$ is the Boltzmann constant, $T$ is the interconnect temperature (in K), $J_s$ is the stress current density, $T_s$ is the stress temperature (in K), $U_1$ is the quantile for the given CDF at product's end of life, $t_{EOL}$. $X_0$, $t_{50}$ and $\sigma$ are the three parameters in the lognormal distribution derived from the stress data. In the special case where $U_1$ approaches negative infinity, the acceleration equation (14) is reduced to,

$$J_{use} = J_s \left\{ \frac{X_0}{t_{EOL}} \right\}^{1/N} \times \exp\left[ \frac{\Delta H}{Nk}\left( \frac{1}{T} - \frac{1}{T_s} \right) \right], \qquad (15)$$

and $J_{use}$ is determined only by $X_0$, in addition to the kinetic parameters of $\Delta H$ and $N$, i.e., $J_{use}$ is independent of $t_{50}$ and $\sigma$ from the stress data.

We have described a procedure for estimating $X_0$ confidence bound in the previous section; for EM reliability evaluation, our objective is to understand how the confidence bounds (or errors) of the estimated values of the parameters from the stress data propagate to the $J_{use}$ projection. In this section, our focus will be on estimation of $J_{use}$ as a function of the three distribution parameters, $X_0$, $t_{50}$ and $\sigma$. One must recognize, however, that the errors in kinetic parameters, $\Delta H$ and $N$, can also have significant effect on the projected $J_{use}$. For the sake of simplicity, we will ignore the effect of these parameters in the present article.

By (15), $J_{use}$ is a monotonic function of $X_0$ for a given set of stress data and product design,

$$J_{use} = \psi(X_0) \qquad (16)$$

In this case, the confidence bound (say for 95%) of $J_{use}$ can be directly computed by replacing the $X_0$ value with its lower 95% confidence bound value,

$$LowerBound[1-\alpha, \psi(X_0)] = \psi[LowerBound(1-\alpha, X_0)] \quad (17)$$

However, for general cases, equation (14) shows that $J_{use}$ is a function of all the three parameters,

$$J_{use} = \psi(t_{50}, \sigma, X_0), \quad (18)$$

and thus errors in all three parameters propagate to $J_{use}$ estimation. To include contributions of all these three parameters to the $J_{use}$ confidence bound, we can re-parametrize any one of these three parameters ($X_0$, $t_{50}$, $\sigma$) in the likelihood function (2) and conduct a similar likelihood ratio testing as that for the $X_0$ confidence bound estimation. If $X_0$ is re-parametrized by $J_{use}$, based on equation (14),

$$X_0(J_{use}, t_{50}, \sigma) = \frac{t_{50}\exp(U_1\sigma) - t_{EOL}(J_x / J_{use})^{-N}}{\exp(U_1\sigma) - 1}, \quad (19)$$

where

$$J_x = J_s \times \exp\left[\frac{\Delta H}{Nk}\left(\frac{1}{T} - \frac{1}{T_s}\right)\right]. \quad (20)$$

The re-parametrized likelihood function is constructed by replacing $X_0$ in equation (2) with $J_{use}$ through equation (19). The steps of $J_{use}$ confidence bound estimation with profile likelihood are the same as for $X_0$. The results are shown in Figure 7. In this particular example of Case A from Figure 2, the lower 95% confidence bound for $J_{use}$ is 1.09MA/cm$^2$, while the estimated nominal $J_{use}$ from equation (14) is 1.74 MA/cm$^2$. The values of parameters in equations (19) and (20) used for these estimations are listed in Table II. These values are for illustration purpose for this specific example, and may not be taken as generic numbers for different interconnect integration and EM structure designs.

One can obtain bias-corrected estimates and confidence bounds for $J_{use}$ by using the same methodology as that described in Procedure A. This correction could be very important when the projected $J_{use}$ is close to the allowed values in product design.

## 6. Discussion

In this paper, the focus has been on the case of $X_0$ with positive values. Mathematically, non-positive values of $X_0$ may be obtained from fitting some of the stress data. Unlike the positive $X_0$, there is no physical meaning for $X_0$ with negative value, considering that EM is a wear out failure mechanism. Graphically, when the early portion of a failure time distribution curve in the probability plot does not bend downwards, it indicates a non-positive $X_0$ in the three-parameter lognormal distribution. This may be an indication of process defects (or non-uniform distribution) of the hardware, further discussions on these special cases are beyond the scope of this paper.

To focus on the methodology of the three-parameter lognormal estimation for EM stress data, this paper also does not attempt to discuss the details of the kinetics aspect of EM lifetime -- in the likelihood re-parametrization process for $J_{use}$ estimation, $\Delta H$ and $N$ were considered as constants. For more accurate lifetime projection, one may also want to explore the propagation of errors in these constants to the projected $J_{use}$ .

## 7. Summary

While three-parameter lognormal distribution can fit the data better and generate more reasonable lifetime projections than two-parameter lognormal, especially for the structures with insufficient redundancy, presence of an additional threshold parameter $X_0$ also makes the data analysis more complicated and challenging. Parameter estimation based on MLE in conjunction with parametric bootstrap procedures for bias correction resulted in meaningful and informative inference for the experimental data. Examples are also given on confidence bound estimation and its propagation to lifetime projections using profile likelihood.

## 8. Acknowledgements

The authors would like to acknowledge the helpful discussions with their BEOL Reliability colleagues, and the management support from Dinesh Badami of IBM semiconductor technology reliability.

## 9. References

[1]. 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.

[2]. J. Gill, et al, " Investigation of Via-Dominated Multi-Modal EM Failure Distributions in Dual Damascene Cu Interconnects with a Discussion of the Statistical Implications", Proc. 40th Annual Reliability Physics Symposium, IEEE, 2002, pp298-304

[3]. E. G. Liniger, et al, "Effects of liner thickness on electromigration lifetime," J. Appl. Phys., 2003, 93(13), pp9576-9582.

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

[5]. B. Li, et al, "Minimum void size and 3-parameter lognormal distribution for EM Failures in Cu Interconnects,", to be published.

[6]. J. R. Black, "Mass transport of Aluminum by momentum exchange with conducting electrons," Proc. 43rd Annual Reliability Physics Symposium, IEEE, 2005, pp1 -6 (Reprinted from 6th Annual Reliability Physics Symposium, 1968).

[7]. JEDEC/FAS Joint Publication, "Foundry process qualification guidelines," JP001, JEDEC, Sept, 2002.

[8]. JEDEC Standard, JESD37, "Standard for lognormal analysis of uncensored data and of singly right-censored data utilizing the Persson and Rootzen method,", Electronic Industries Association, Oct, 1992.
[9]. R. G. Filippi, et al; "Paradoxical predictions and minimum failure time in electromigration", Appl. Phys. Lett., 66(16), 1995,pP1897-1899
[10]. J. L. Devore, Probability and Statistics, Pp265-274, 5$^{th}$ edition, Duxbury, 2000.
[11]. W. Nelson, Applied Life Data Analysis, pp544, John Wiley & Sons, 1982.
[12]. N. L. Johnson, S. Kotz and N. Balakrishnan "Continuous Univariate Distributions", Vol.1, Wiley, 1994.
[13]. W. Q. Meeker and L. A. Escobar "Statistical Methods for Reliability Data", Wiley, 1998.

**Table I. Goodness of Fit Comparison between 2- and 3-Parameter Lognormal**

| Technology Node | 6nm (thin wire) | 65nm (fat wire) | 90nm | 180nm |
|---|---|---|---|---|
| Symbol in Fig. 2 | A | B | C | D |
| Via size | (0.1x0.1) $\mu m^2$ | (0.2x0.2) $\mu m^2$ | (0.14x0.14) $\mu m^2$ | (0.28x0.28) $\mu m^2$ |
| Line width | 0.3μm | 0.6μm | 0.28μm | 0.76μm |
| Goodness of fit 3-parameter | 0.003 | 0.013 | 0.017 | 0.006 |
| Goodness of fit 2-parameter | 0.009 | 0.018 | 0.036 | 0.015 |

* Goodness of fit here is defined as the mean of the squared residuals.

**Table II. Parameter Values Used for $J_{use}$ Projection**

| Parameter | Value | Units |
|---|---|---|
| $t_{EOL}$ | 110,000 | hours |
| U1 | -6.7 | |
| $J_s$ | 2.5 | $MA/cm^2$ |
| T | 373 | K |
| $T_s$ | 573 | K |
| n | 1.1 | |
| $\Delta H$ | 0.9 | eV |
| k | 0.0000862 | eV/K |

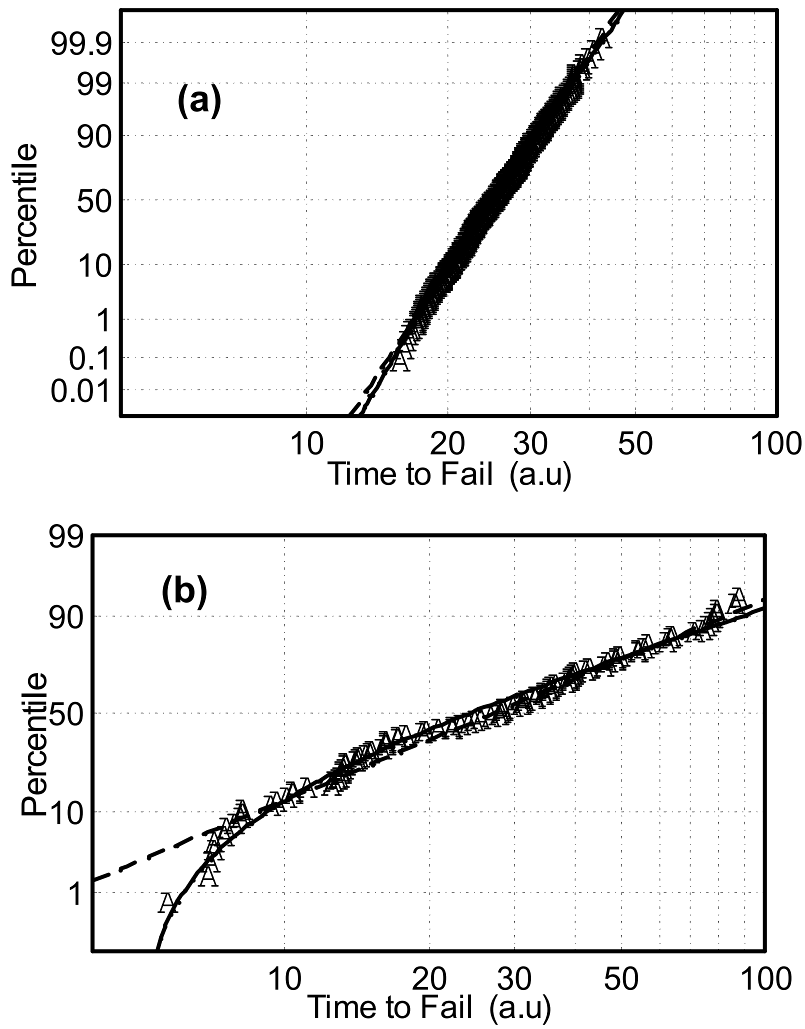



Figure 1. Effect of shape factor (σ) on the difference between two-parameter and three-parameter lognormal fittings. All data were generated by Monte Carlo simulations, with inputs: $t_{50} = 25$, $X_0 = 4$; $\sigma = 0.2$ for (a) and $\sigma = 1.0$ for (b). The solid line is for three-parameter, and broken line is for two-parameter lognormal fittings.

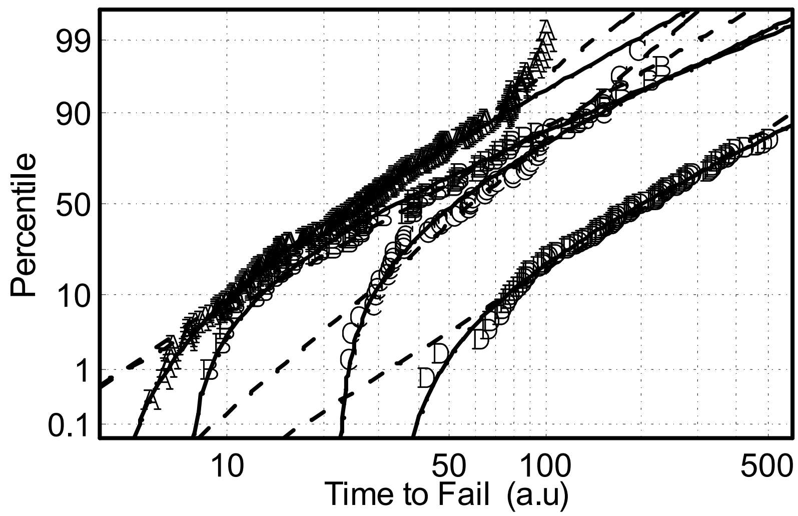
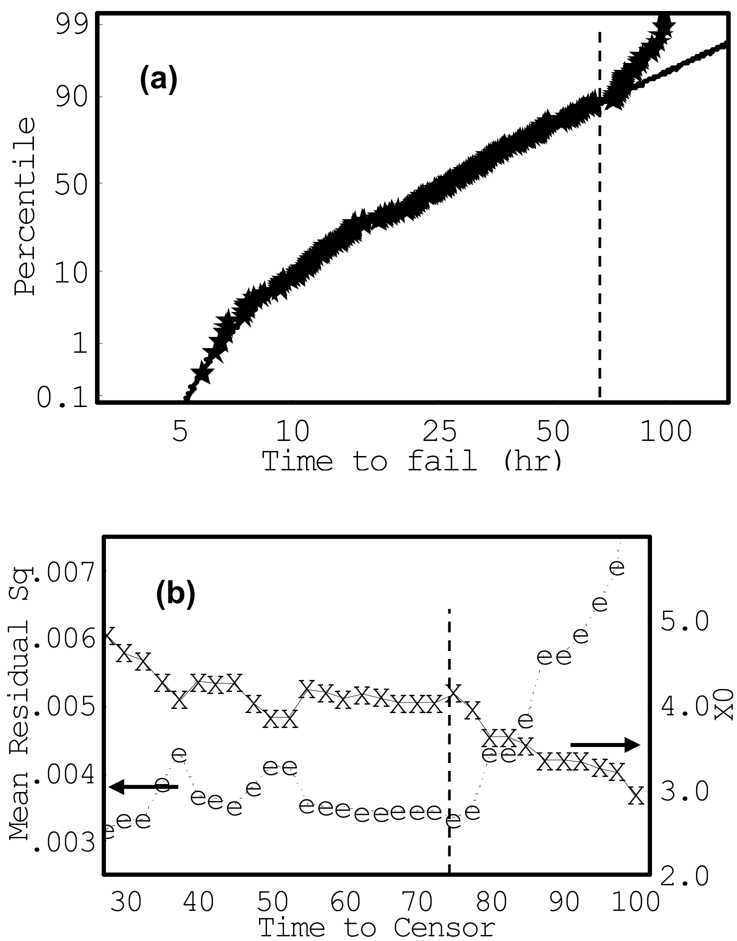



Figure 2. Examples of EM failure time distributions of structures with a single via connecting a wide line below from different technology nodes. All stresses were conducted at 300 $^{o}$C, 2.5MA/cm$^{2}$. Solid lines are for three-parameter, and broken lines are for two-parameter lognormal fitting. Symbols are defined in Table 1.



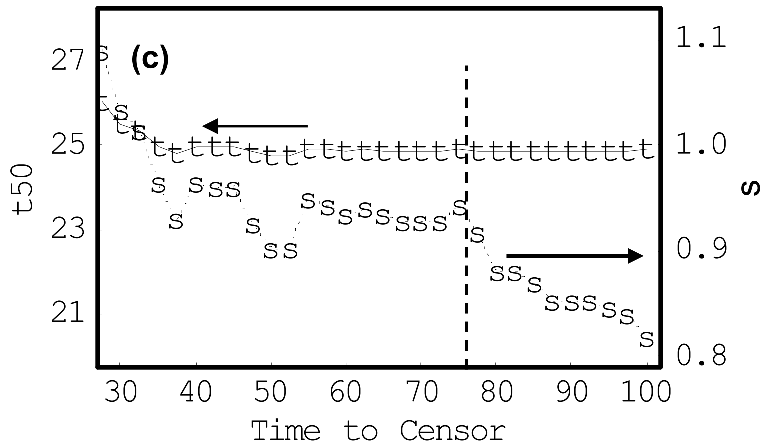



Figure 3. Fitting parameter variation with right data censoring on the experimental data of Case A from Figure 2. (a) three-parameter fitting on the EM data; (b) estimated $X_0$ (symbol "X") and goodness of fit (symbol "e") variation and (c) $t_{50}$ (symbol "t") and σ (symbol "s") variation.

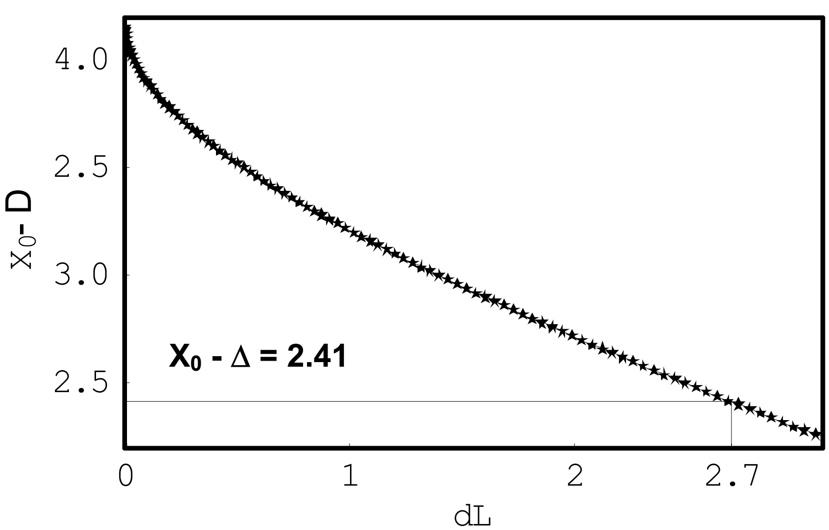



Figure 4. $X_0$ confidence bound estimation using profile likelihood. ($X_0 - \Delta$) is the fixed $X_0$ values used in the constrained MLE, and $dL$ is defined in equation (9)

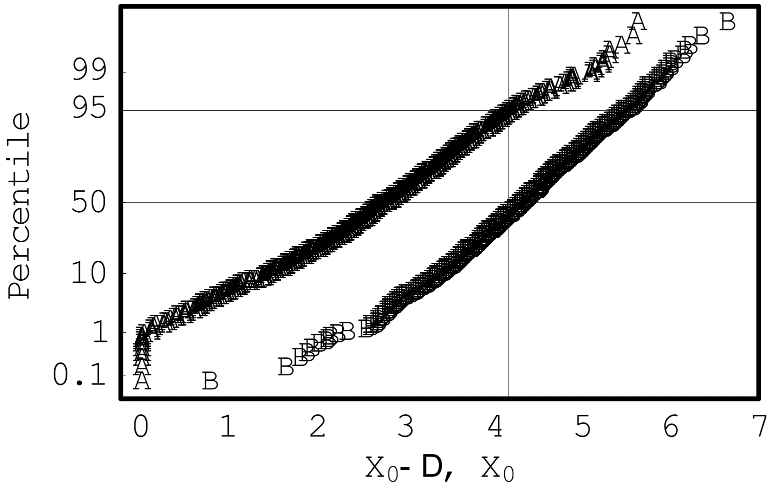



Figure 5. Monte Carlo simulation results of 1000 runs on $X_0$ (symbol B) and $X_0$-Δ (symbol A) estimation. Each run has a sample size of 240 points. The input three-parameter lognormal parameters are $X_0$ = 4.16, $t_{50}$ = 24.90, σ = 0.94.

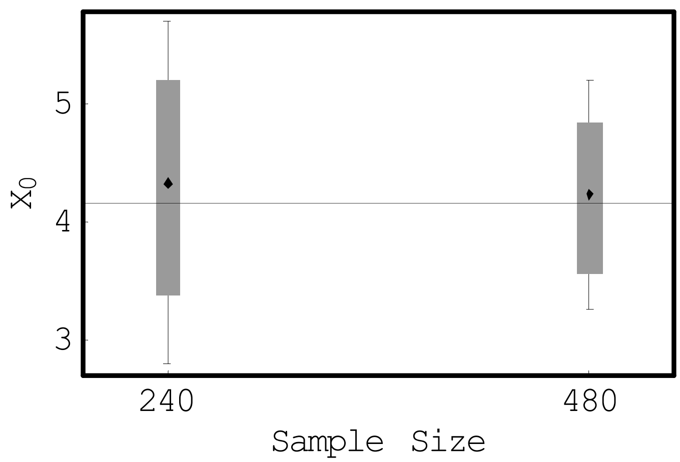



Figure 6. Effect of sample size on the upward bias of the estimated $X_0$ values from 1000 runs of Monte Carlo simulation. The black dot is the median value of the estimated $X_0$ and the horizontal line is the expected $X_0$ value (the input value for the simulation), the gray bar covers10 – 90 percentiles, the vertical line bar covers 2.5 – 97.5 percentiles.

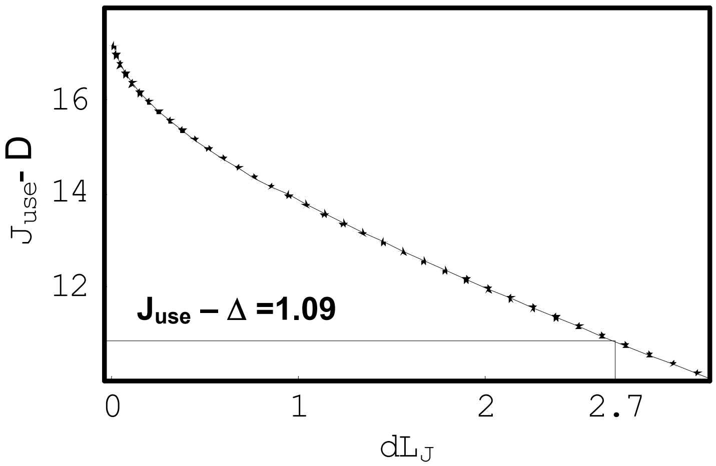


Figure 7. The results of $J_{use}$ confidence bound estimation with re-parametrized profile likelihood. $dL_J$ is similar to dL, but with the likelihood function (2) re-parametrized to $J_{use}$ using equation (19).