

DTIC FILE COPY



RADC-TR-90-93
Final Technical Report
May 1990

VLSI DESIGN FOR RELIABILITY

University of Illinois

Ibrahim N. Hajj, Farid N. Najm, Ping Yang

AD-A224 264

DTIC
ELECTE
JUL 10 1990
S B D

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

Rome Air Development Center
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700

90 07 10 025

This report has been reviewed by the RADC Public Affairs Division (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-90-93 has been reviewed and is approved for publication.

APPROVED: *Mark T. Pronobis*

MARK T. PRONOBIS
Project Engineer

APPROVED: *John J. Bart*

JOHN J. BART
Technical Director
Directorate of Reliability & Compatibility

FOR THE COMMANDER:

James W. Hyde III

JAMES W. HYDE III.
Directorate of Plans & Programs

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (RBRP) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing the instructions, searching existing data sources, gathering and maintaining the data needed, and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Information and Regulatory Affairs, Office of Management and Budget, Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE May 1990	3. REPORT TYPE AND DATES COVERED Final Sep 89 to Nov 89	
4. TITLE AND SUBTITLE VLSI DESIGN FOR RELIABILITY			5. FUNDING NUMBERS C - F30602-88-D-0028, Task N-9-0125 PE - 62702F PR - 2338 TA - 01 WU - PX	
6. AUTHOR(S) Ibrahim N. Hajj, Farid N. Najm, Ping Yang			8. PERFORMING ORGANIZATION REPORT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Illinois Coordinated Science Lab 1101 W. Springfield Ave. Urbana IL 61801			10. SPONSORING/MONITORING AGENCY REPORT NUMBER RADC-TR-90-93	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Rome Air Development Center (RBRP) Griffiss AFB NY 13441-5700			11. SUPPLEMENTARY NOTES RADC Project Engineer: Mark T. Pronchbis/RBRP/(315) 330-4995	
12a. DISTRIBUTION/AVAILABILITY STATEMENT proved for public release; distribution unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This report contains the results of supplementary work done related to the reliability analysis of Application Specific Very Large Scale Integrated (ASIC VLSI) CMOS circuits. The major work is currently being carried out under Task N-9-5716. The main goal of both tasks is to determine the electromigration susceptibility of VLSI circuits. Electromigration is a major reliability problem caused by the transport of atoms in a metal line due to the electron flow. Under persistent current stress, electromigration can cause deformations of the metal lines which may result in shorts or open circuits. The failure rate due to electromigration depends on the current density in the metal lines and is usually expressed as a median-time-to-failure (MTF). This work focuses on the electromigration problem in the power and ground busses. To estimate the bus MTF, an estimate of the current waveform in each branch of the				
14. SUBJECT TERMS Reliability, probabilistic simulation, variance waveform, electromigration, VLSI CMOS circuits			15. NUMBER OF PAGES 28	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED			18. PRICE CODE	
18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED		19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED		20. LIMITATION OF ABSTRACT UL

٧

lines

$$L(KR) \leftarrow$$

EVALUATION

This technical report describes a technique for including the current variance waveform in the calculation for the median-time-to-failure (MTF) for metal power and ground busses due to electromigration in VLSI integrated circuits. Previously described work details the use of probabilistic simulations in determining the expected current waveforms in the branches of the busses. This work establishes the relationship between the MTF and the expected and variance waveforms in the bus. Several examples are presented which compares the current waveforms obtained by using SPICE simulations and waveforms obtained using probabilistic simulation with and without the variance contribution. The main advantage of this approach is the ability to handle large circuits by replacing an exponential number of deterministic simulation runs with a single probabilistic simulation run.

Mark T. Pronobis

MARK T. PRONOBIS
Reliability Physics Branch
Microelectronics Reliability Division



Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

1. Introduction

Under this task, we performed supplementary work related to the reliability analysis of Application Specific Very Large Scale integrated (ASIC VLSI) CMOS circuits. The major work is currently being carried out under Task N-9-5716. The main goal of both tasks is to determine the electromigration susceptibility of VLSI circuits. Electromigration is a major reliability problem caused by the transport of atoms in a metal line due to the electron flow. Under persistent current stress, electromigration can cause deformations of the metal lines which may result in shorts or open circuits. The failure rate due to electromigration depends on the current density in the metal lines and is usually expressed as a median-time-to-failure (MTF).

In this work we focus our attention on the electromigration problem in the power and ground busses. To estimate the bus MTF, an estimate of the current waveform in each branch of the bus is required. In general, the MTF is dependent on the *shape* of the current waveform, and not simply on its time-average. However, a very large number of such waveform shapes are possible, depending on what inputs are applied to the circuit. This is especially true for CMOS circuits, which draw current only during switching. The correct current waveform to be used for MTF estimation is one that combines (in some sense) the effects of all possible logic input waveforms. If we consider the set of logic waveforms allowed at the circuit inputs as a *probability space*, the current in any branch of the bus becomes a *stochastic process*. The *expected current waveform* is then the *mean waveform* (not a time-average) of this process. This is a waveform whose value at a given time is the weighted average of all possible current values at that time. The resulting methodology is what we call a *probabilistic simulation* of the circuit. We have implemented this approach in CREST [1] which uses *statistical* information about the inputs to directly derive the expected current waveforms.

In this task, we have established that the MTF estimate not only depends on the expected current waveform, but also on the current *variance waveform*, especially in regions where the current density is high. This report presents our findings and is organized as follows. In the next section, we establish the relationship between the MTF and the expected and variance current waveforms in the bus. In Section 3, we describe how an estimate of the variance current waveform is derived at the terminals of CMOS gates. In Section 4, we explain the approach to be used in estimating the variance current waveform anywhere within a bus given the variance current waveforms at the contacts to the bus. Implementation issues and results are presented in Section 5.

2. Stochastic Current Waveforms and the MTF

Consider a metal line of uniform width and thickness carrying a *constant* current. The relationship between the MTF, t_{50} , due to electromigration in the metal line and the current density, j , has been extensively studied, and shown to be a complex nonlinear function [2], as shown in Fig. 1. We will consider the MTF to be $t_{50} \propto 1/f(j)$ where j is in A/cm^2 , and f is a dimensionless nonlinear function, whose plot is shown in Fig. 2, which was derived from Fig. 1.

If a metal line carries a *varying* current, of density $j(t)$, then the MTF is $t_{50} \propto 1/J_{eff}$, where J_{eff} depends both on f and on the waveform shape of $j(t)$. It has been suggested [3] that, if the waveform is periodic with period T and consists of a train of pulses $k = 1, \dots, m$ of heights j_k and duration t_k , then :

$$J_{eff} = \sum_{k=1}^m \frac{t_k}{T} f(j_k). \quad (1)$$

For a general periodic waveform, we take the summation to the limit and write :

$$J_{eff} = \frac{1}{T} \int_0^T f(j) dt. \quad (2)$$

If the current waveform is not periodic, then better estimates of J_{eff} are obtained by using larger values of T so that more features of the waveform are included. Therefore one can write :

$$J_{eff} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(j) dt. \quad (3)$$

Now suppose that the current waveform is *stochastic*, i.e., it is a stochastic process $j(t)$, that represents a family of deterministic (real) current waveforms $j_k(t)$, with associated probabilities P_k , $k = 1, \dots, N$, over the (finite) interval $[0, t_0]$. Based on this information, we can build a (non-stochastic) current waveform $j(t)$, over $[0, T]$ as $T \rightarrow \infty$, that is indicative of the current during typical operation as follows. Consider a random sequence of the waveforms $j_k(t)$, each being shifted in time, spanning an interval of length t_0 , and occurring with its assigned probability P_k , as shown in Fig. 3. Let $n_k(T)$ be the (integer) number of occurrences of the waveform $j_k(t)$ in $[0, T]$, and let $n_T = \lfloor T/t_0 \rfloor$. If J_k , $k = 1, \dots, N$ are defined as follows :

$$J_k \triangleq \frac{1}{t_0} \int_0^{t_0} f(j_k) dt, \quad (4)$$

then :

$$J_{\text{eff}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(j) dt = \lim_{n_T \rightarrow \infty} \sum_{k=1}^N J_k \frac{n_k(T)}{n_T} = \sum_{k=1}^N J_k \lim_{n_T \rightarrow \infty} \left[\frac{n_k(T)}{n_T} \right]. \quad (5)$$

By the law of large numbers [4], $\lim_{n_T \rightarrow \infty} [n_k(T)/n_T] = P_k$, which leads to :

$$J_{\text{eff}} = \sum_{k=1}^N \left[\frac{1}{t_0} \int_0^{t_0} f(j_k) dt \right] P_k = \frac{1}{t_0} \int_0^{t_0} \left[\sum_{k=1}^N f(j_k) P_k \right] dt \quad (6)$$

and, finally :

$$J_{\text{eff}} = \frac{1}{t_0} \int_0^{t_0} E[f(j)] dt \quad (7)$$

where $E[]$ denotes the expected value operator. This is an important result; it says that *the MTF due to a stochastic current depends only on the expected waveform of a nonlinear function of the current.*

Since f is nonlinear, $E[f(j)]$ is not easy to evaluate. At low current values, where f is linear (Fig. 2), $E[f(j)] = f(E[j])$. If this is substituted in (7) and compared with (2) it shows that, when f is linear, the expected current waveform $E[j]$ derived in [1] may itself be used as the current waveform $j(t)$ in (2) for MTF estimation. This establishes the importance of the expected current waveform for electromigration failure analysis. In general, f is nonlinear, and a generalized approach will be developed below.

At any time t , the process $j(t)$ can be thought of as a random variable j with mean $\eta_j \triangleq E[j]$, and variance $\sigma_j^2 \triangleq E[(j - \eta_j)^2]$. In general, the p^{th} moment of j is $\mu_{j,p} \triangleq E[(j - \eta_j)^p]$. To estimate the mean of $f(j)$, $E[f(j)]$, we use a Taylor series expansion of f , which leads to :

$$E[f(j)] \approx f(\eta_j) + f''(\eta_j) \frac{\sigma_j^2}{2} + \dots + f^{(p)}(\eta_j) \frac{\mu_{j,p}}{p!}. \quad (8)$$

It is evident that, when f is linear, (8) reduces to :

$$E[f(j)] = f(E[j]), \quad (9)$$

as observed above. Hence using the expected current waveform as an actual current waveform for MTF estimation based on (2) amounts to making a first-order approximation in (8). Naturally, higher order approximations would lead to better results. In particular, if f is approximated by a quadratic in the neighborhood of η_j , then :

$$E[f(j)] \approx f(\eta_j) + f''(\eta_j) \frac{\sigma_j^2}{2}. \quad (10)$$

This second-order approximation becomes exact if $f(j)$ is represented by the straight lines corresponding to j^1 and j^2 in Fig. 2. It is more accurate than (9) since it covers a wider range of currents. As a result, *equations (10) and (7) offer a new, more accurate technique for computing the MTF*. In order to make use of this technique, we need to derive the variance of the current waveform in addition to its expected value. As pointed out in the introduction, the estimation of the expected current waveform has already been described in our previous work [1]; the next section will discuss the derivation of the variance.

3. Derivation of the Gate Variance Waveforms

Figure 4 shows a generic CMOS gate structure which we will use in our derivation. The variance waveforms for the gate total and output currents will be modeled by triangular pulses $V[i_{tot}(t)]$ and $V[i(t)]$, respectively, with peak values of $V[I_{tot}]$ and $V[I]$. If an event occurs at the gate input at time t , then we denote by t^- and t^+ the instances of time immediately before and after the event, respectively. Focusing for now on the output current pulse, its variance waveform starts with a peak of $V[I] = V[i(t^+)]$ at time t and decays linearly

to zero at time $t + \tau$. Since $V[I] = E[I^2] - E[I]^2$ [4], and since CREST already derives the expected pulse peak ($E[I]$), we will concentrate here on the derivation of $E[I^2]$.

Let $i_p = i_{p1} + i_{p2}$ and $i_n = i_{n1} + i_{n2}$. It is easy to verify that $i_{p1} = i_p \times C_n / (C_p + C_n)$, and $i_{n1} = i_n \times C_p / (C_p + C_n)$. Therefore :

$$E[i^2(t)] = E[i_p^2(t)] \left(\frac{C_n}{C_p + C_n} \right)^2 + E[i_n^2(t)] \left(\frac{C_p}{C_p + C_n} \right)^2$$

The term containing $E[i_p(t)i_n(t)]$ is omitted (it is zero) since at least one of the charging currents is zero at any given time. In particular, the value at the peak is :

$$E[I^2(t)] = E[I_p^2(t)] \times \left(\frac{C_n}{C_p + C_n} \right)^2 + E[I_n^2(t)] \times \left(\frac{C_p}{C_p + C_n} \right)^2$$

The values of $E[I_p^2]$ and $E[I_n^2]$ are derived as follows. For $E[I_p^2]$, consider the p-part of the gate, and let every transistor T_k be represented by a switch of on-conductance $g_{on,k}$ [5]. Based on this switch-network model of the p-block, let $G_p(t)$ be the random conductance between the output node and V_{dd} . G_p is a function of the individual transistor random conductances g_k , where g_k is 0 if the transistor is off and $g_{on,k}$ if it is on. If an event occurs at the gate at time t , then the value of $G_p(t^+)$ and the previous state of the output node, $V_o(t^-)$, will determine I_p . Formally, we have $E[I_p^2] = E[(V_{dd} - V_o(t^-))^2 \times G_p^2(t^+)]$, which becomes :

$$E[I_p^2] = V_{dd}^2 \times E[G_p^2(t^+) \mid G_p(t^-) = 0] \times P(G_p(t^-) = 0)$$

where $P(A)$ is the probability of the event A , and $E[A \mid B]$ denotes the *conditional expected value* [4] of A given B . The formula is correct because if $G_p(t^-) = 0$ ($\neq 0$) then $V_o(t^-) = 0$ (V_{dd}). Similarly for the n-part of the gate, we get :

$$E[I_n^2] = V_{dd}^2 \times E[G_n^2(t^+) \mid G_n(t^-) = 0] \times P(G_n(t^-) = 0)$$

To derive the conditional expectations, consider a graph representation of the p-block (or n-block), where every edge in the graph is labeled with $E[g_k^2(t^+) \mid G_p(t^-) = 0]$, $E[g_k(t^+) \mid G_p(t^-) = 0]$, and the gate node probabilities of its corresponding transistors. The details of how these quantities can be derived for every transistor can be found in [5].

Then perform a *graph reduction* operation, which, simply stated, involves a number of series/parallel combinations and node eliminations that reduce the graph to a single edge, whose labels are the required statistics $E[G_p^2(t^+) | G_p(t^-) = 0]$ and $E[G_p(t^+) | G_p(t^-) = 0]$. Similarly for the n-block.

Having found the peak $V[I] = E[I^2] - E[I]^2$ for the output current, the time span τ will be found by first solving for the area under the $V[i(t)]$ pulse. Notice that, if $i(t)$ is a triangular pulse of height I and area q , then :

$$\int_0^\infty i^2(t)dt = \frac{2}{3}Iq$$

In this case, q is equal to the charge delivered to (or from) the output node capacitors. From this it follows that :

$$\int_0^\infty E[i^2(t)]dt = \frac{2}{3}E[Iq], \text{ and } \int_0^\infty E[i(t)]^2dt = \frac{2}{3}E[I]E[q].$$

The second equation follows since $E[i(t)]$ is a triangular pulse of height $E[I]$ and area $E[q]$. Therefore, the variance pulse has an area :

$$\frac{V[I] \times \tau}{2} = \int_0^\infty V[i(t)]dt = \frac{2}{3}(E[Iq] - E[I]E[q]).$$

The value of $E[Iq]$ can be written as :

$$E[Iq] = E[(I_{p1} + I_{n1}) \times (q_{p1} + q_{n1})].$$

where I_{p1} (I_{n1}) is the peak of $i_{p1}(t)$ ($i_{n1}(t)$), and q_{p1} (q_{n1}) is the charge delivered by $i_{p1}(t)$ ($i_{n1}(t)$). When the p-block is conducting, the n-block is not conducting, and $i_{p1}(t) \neq 0$, $i_{n1}(t) = 0$, $q_{p1} \approx V_{dd}C_n$, and $q_{n1} \approx 0$; while when the n-block is conducting, the p-block is not, and $i_{n1}(t) \neq 0$, $i_{p1}(t) = 0$, $q_{n1} \approx V_{dd}C_p$, and $q_{p1} \approx 0$. It follows then that:

$$E[Iq] = \frac{V_{dd}C_n^2}{C_p + C_n}E[I_p] + \frac{V_{dd}C_p^2}{C_p + C_n}E[I_n].$$

The time span of the gate output current variance pulse is, therefore :

$$\tau = \frac{4}{3} \left(\frac{E[Iq] - E[I]E[q]}{V[I]} \right).$$

If $i_{tot}(t)$ is the total gate current, then :

$$\int_0^\infty V[i_{tot}]dt = \frac{2}{3} (E[I_{tot}q_{tot}] - E[I_{tot}]E[q_{tot}]).$$

Unfortunately, $E[I_{tot}q_{tot}]$ does not have a simple expression as was found for $E[Iq]$ above. We have chosen to use a conservative estimate based on the following assumption : whenever a node in the p-block (n-block) is charged to V_{dd} (V_{ss}), then every other node in the p-block (n-block) is also charged to V_{dd} (V_{ss}). This assumption is true for simple gates, and overestimates the current charge product in more complex cases. Based on this assumption, one can show [5] that :

$$E[I_{tot}q_{tot}] \approx \frac{E[q_{p,tot}]}{E[q_p]} \frac{Q_p C_n E[I_p]}{C_p + C_n} + \frac{E[q_{n,tot}]}{E[q_n]} \frac{Q_n C_p E[I_n]}{C_p + C_n},$$

where $E[q_p]$ and $E[q_n]$ are available as equations (9) & (10) in [6], $E[q_{p,tot}]$ and $E[q_{n,tot}]$ are, respectively, the first and second summations in equation (7) in [6], and

$$Q_p = \sum_{i \in P \text{ block}} V_{dd} C_{in}, \quad Q_n = \sum_{i \in N \text{ block}} V_{dd} C_{ip}.$$

As was assumed for the expected current pulse [5, 6], we let the time span of the gate total current variance pulse be equal to that derived for the gate output current, therefore :

$$V[I_{tot}] = \left(\frac{E[I_{tot}q_{tot}] - E[I_{tot}]E[q_{tot}]}{E[Iq] - E[I]E[q]} \right) \times V[I].$$

4. Estimating the Variance Current Waveforms in the Bus

Since the current density $j(t)$ in any branch of the power or ground bus is directly proportional to the current $i(t)$ in that branch, then, to simplify the presentation, we will discuss the derivation of σ_i^2 rather than σ_j^2 . The variance of the current waveform is a time function $\sigma_i^2(t)$ which we will refer to as the *variance waveform*. Furthermore, we will discuss the power bus only since the ground bus analysis is similar.

The current in a branch of the bus, $i(t)$, is a function of the currents being drawn off the bus contacts, $i_j(t), j = 1, \dots, n$. Each of these is, in turn, simply the sum of the individual gate currents tied to each contact :

$$i_j(t) = i_{j1}(t) + \dots + i_{jk}(t). \quad (11)$$

Thus, in the framework of our probabilistic simulation technique, the process of deriving the variance waveforms consists of three steps :

- 1- Using the statistics of the signals at the inputs to each logic gate, derive the variance waveform for its current.
- 2- Combine these variance waveforms at each contact point to derive the variances of the contact currents.
- 3- Using the bus topology, and the variances of the contact currents, derive the variances of the bus branch currents.

Step 1 has been described in Section 3 above. The other two steps will be described below.

The critical issue is the correlation between the different current waveforms. Since such correlation is too expensive to derive for VLSI circuits, we will occasionally be making conservative approximations to simplify the problem. Our experience with the probabilistic simulation approach suggests that neglecting the correlation between different current waveforms gives good results in most cases.

Based on this premise, we assume that the gate currents tied to the same contact are *uncorrelated*. This immediately provides a simple solution for step 2, using (11), as follows :

$$\sigma_{ij}^2(t) = \sigma_{i,j_1}^2(t) + \dots + \sigma_{i,j_k}^2(t). \quad (12)$$

The remainder of this section will be devoted to the more difficult task of solving step 3, i.e., deriving the bus current variance waveforms from those of the contact currents.

The metal bus can be modeled as a multi-input multi-output, causal, *linear*, time-invariant, (LTI) system with causal inputs x_j and outputs y_i . The inputs $x_j(t)$, $j = 1, \dots, n$ represent the contact currents, and carry the stochastic processes $i_j(t)$ of known variance waveforms $\sigma_{i_j}^2(t)$. The outputs $y_i(t)$, $i = 1, \dots, m$ represent the bus branch currents at which the variance waveforms, $\sigma_{y_i}^2(t)$, are required. Let $h_{ij}(t)$ be the impulse response function relating $y_i(t)$ to $x_j(t)$:

$$y_i(t) = \sum_{j=1}^n h_{ij}(t) * x_j(t), \quad i = 1, \dots, m \quad (13)$$

where "*" denotes the *convolution* operation.

It is well known (see [4], page 209) that the variances of the system inputs are not enough to derive the variances of its outputs. The *auto-correlation* of each input, $R_{x_j x_j}(t_1, t_2) \triangleq E[x_j(t_1)x_j(t_2)]$, is also required. Since the input processes are not *wide-sense stationary* [4], an exact analytical solution can be quite complex, even if the auto-correlation were known. Therefore, as is often necessary, we will make certain simplifying assumptions about the structure of $R_{x_j x_j}$.

We will assume that the correlation between $x_j(t)$ and $x_j(t + \tau)$ goes to zero as $\tau \rightarrow \infty$. In terms of the *auto-covariance*, $C_{x_j x_j}(t_1, t_2) \triangleq R_{x_j x_j}(t_1, t_2) - \eta_{x_j}(t_1)\eta_{x_j}(t_2)$, this will be formulated as :

$$C_{x_j x_j}(t_1, t_1) = \sigma_{x_j}^2(t_1), \text{ and } C_{x_j x_j}(t_1, t_2) = 0 \text{ for } |t_1 - t_2| \geq T, \quad (14)$$

where T is a (typically small) time interval.

Consider the *discrete time* system obtained by sampling, with period T , the continuous time system defined by (13). If $x_j[k] \triangleq x_j(kT)$ are the discrete processes at the inputs, and $y_i[k] \triangleq y_i(kT)$ are the discrete output processes, then :

$$y_i[k] = \sum_{j=1}^n h_{ij}^{(d)}[k] * x_j[k], \quad i = 1, \dots, m \quad (15)$$

where $h_{ij}^{(d)}[k]$ is the discrete impulse response function relating $y_i[k]$ to $x_j[k]$. As shown below, the discretized output variance waveforms can be derived irrespective of the shape of $C_{x_j x_j}(t_1, t_2)$ for $|t_1 - t_2| < T$. The continuous variance waveforms can then be obtained by interpolation. Strictly speaking, therefore, the sampling period T should be small : $1/T$ should be larger than the largest frequency component of the inputs. However, since fine waveform details are not of paramount importance in this work, we need only restrict T to be small enough so that waveform features in so small an interval are inconsequential.

To simplify the notation, define $y_{ij}[k] \triangleq h_{ij}^{(d)}[k] * x_j[k]$. Furthermore, as pointed out above, we will neglect the correlation between the contact currents. Hence the x_j inputs are uncorrelated, and :

$$\sigma_{y_i}^2[k] = \sum_{j=1}^n \sigma_{y_{ij}}^2[k], \quad i = 1, \dots, m. \quad (16)$$

We have thus reduced the problem to analyzing a single-input single-output discrete LTI system :

$$y_{ij}[k] = h_{ij}^{(d)}[k] * x_j[k] = \sum_{\kappa=0}^{\infty} h_{ij}^{(d)}[\kappa] x_j[k - \kappa]. \quad (17)$$

Let $\tilde{x}_j[k] \triangleq x_j[k] - \eta_{x_j}[k]$ and $\tilde{y}_{ij}[k] \triangleq y_{ij}[k] - \eta_{y_{ij}}[k]$. Then $\sigma_{x_j}^2[k] = E[\tilde{y}_{ij}[k]^2]$ and $\tilde{y}_{ij}[k] = h_{ij}^{(d)}[k] * \tilde{x}_j[k]$, hence :

$$\sigma_{y_{ij}}^2[k] = E \left[\left(\sum_{\kappa=0}^{\infty} h_{ij}^{(d)}[\kappa] \tilde{x}_j[k - \kappa] \right)^2 \right] = \sum_{\kappa_1=0}^{\infty} h_{ij}^{(d)}[\kappa_1] \sum_{\kappa_2=0}^{\infty} h_{ij}^{(d)}[\kappa_2] E[\tilde{x}_j[k - \kappa_1] \tilde{x}_j[k - \kappa_2]]. \quad (18)$$

Furthermore, it is easy to see that $E[\tilde{x}_j[k_1] \tilde{x}_j[k_2]] = C_{x_j x_j}(k_1, k_2)$, which, using (14), gives :

$$\sigma_{y_{ij}}^2[k] = \sum_{\kappa=0}^{\infty} |h_{ij}^{(d)}[\kappa]|^2 \sigma_{x_j}^2[k - \kappa] = |h_{ij}^{(d)}[k]|^2 * \sigma_{x_j}^2[k]. \quad (19)$$

And, finally, the variance waveforms for the system outputs are, using (16) :

$$\sigma_{y_i}^2[k] = \sum_{j=1}^n |h_{ij}^{(d)}[k]|^2 * \sigma_{x_j}^2[k], \quad i = 1, \dots, m. \quad (20)$$

In other words, *the variances of the system outputs (bus branch currents) can be obtained from the convolution of the variances of its inputs (contact currents) with the squares of its discrete impulse response functions*. This discrete convolution can be easily performed once the discrete impulse response functions are found. Of course the summation need not be taken to infinity, and may be conveniently truncated after $|h_{ij}^{(d)}[\kappa]|$ is less than some small value. To obtain the discrete impulse response functions, note that if a unit-step input current is applied at contact j , with all other contact currents held at zero, and if the resulting outputs $y_i(t)$ are monitored, then :

$$h_{ij}^{(d)}[k] = y_i(kT) - y_i((k-1)T) = \int_{(k-1)T}^{kT} h_{ij}(\tau) d\tau, \quad i = 1, \dots, m. \quad (21)$$

This suggests two methods for deriving $h_{ij}^{(d)}[k]$. The first uses a simulation program such as SPICE to simulate the bus with unit-step input currents applied at each contact (one at a time), while monitoring the bus branch currents. This gives the mn functions $h_{ij}^{(d)}[k]$ using (21). Another (approximate) method would be to make use of the second equality in (21) .

if the continuous impulse response functions are approximated using some RC time-constant analysis of the bus, then the discrete impulse response functions can be obtained from them.

For very large chips, it may be prohibitively expensive to perform the required convolutions. One can simplify the calculations by making an additional assumption as follows. If the bus is known to be "fast", i.e., if $h_{ij}^{(d)}[k]$ dies down faster than changes in $\sigma_{x_j}^2[k]$, then (19) reduces to :

$$\sigma_{y_{ij}}^2[k] \approx \sigma_{x_j}^2[k] \sum_{\kappa=0}^{\infty} |h_{ij}^{(d)}[\kappa]|^2. \quad (22)$$

So the convolutions in (20) can be replaced by simple multiplications, and the constants $\sum_{\kappa=0}^{\infty} |h_{ij}^{(d)}[\kappa]|^2$ can be derived in a pre-processing step from the impulse response functions and stored in a single $m \times n$ constant matrix.

If the chip is too big to even derive $h_{ij}^{(d)}[k]$, then one further simplification can be made as follows. If $h_{ij}^{(d)}[k]$ dies down faster than changes in $x_j[k]$ then (17) reduces to $y_{ij}[k] = x_j[k] \sum_{\kappa=0}^{\infty} h_{ij}^{(d)}[\kappa]$, and so :

$$\sigma_{y_{ij}}^2[k] \approx \sigma_{x_j}^2[k] \left(\sum_{\kappa=0}^{\infty} h_{ij}^{(d)}[\kappa] \right)^2. \quad (23)$$

The constants $(\sum_{\kappa=0}^{\infty} h_{ij}^{(d)}[\kappa])^2$ can be very easily obtained as follows. Note that $\sum_{\kappa=0}^{\infty} h_{ij}^{(d)}[\kappa]$ is the *steady state* current in branch i in response to a unit-step input current at contact j , with all other contact currents held at zero. If the bus is modeled as a resistive network, then the steady state node voltages in response to such inputs are the entries of the driving point impedance matrix. So if the node-admittance matrix is built by simple inspection of the bus and then inverted to produce the driving point impedance matrix, the steady state currents are immediately available.

5. Implementation and Results

The variance calculation technique outlined above has been implemented in CREST. We present below the results of CREST runs on a variety of circuits, showing both waveform comparisons and timing performance.

We start out with a simple example, a 2-input CMOS NAND gate. The variance waveform comparison between CREST and SPICE is shown in Fig. 5. The SPICE waveform is

derived by running SPICE on the NAND gate for all possible logical transitions at its inputs, deriving the expected current waveform by doing a time-point averaging of the results, and then using that to find the variance as the time-point average of $(i - E[i])^2$. Since the object of this research is to handle very large chips, and since elect. migration models for ac waveforms are still controversial, it makes little sense to shoot for perfect accuracy in the current waveforms. It is more important to be able to derive in a very short time a waveform that matches the peak and general shape of the SPICE waveform.

Another single-gate comparison is shown in Fig. 6 for a CMOS complex gate. The comparisons for two larger circuits are shown in Fig. 7 (for an XOR circuit) and Fig. 8 (for a 54-MOSFET 2-bit ripple adder circuit).

Our final example is a much bigger and more complex circuit; it is a 64S-MOSFET 4-bit parallel multiplier. This circuit is too big to make the 2^8 required SPICE simulations. We will, therefore, show two different CREST runs to demonstrate that the variance computation works well even when the heuristics introduced for handling large circuits in [1] are used. In Fig. 9 we compare a full accuracy CREST run and a heuristic CREST run in which all internal nodes of the multiplier were assumed independent. The excellent agreement reaffirms the conclusion made in [1] that as a circuit becomes larger, the correlation between its internal nodes may be safely neglected.

We next examine the importance of the variance waveform for MTF estimation. The expected and variance waveforms combine to provide a J_{eff} dc current density value to represent the ac current waveform for MTF estimation as given in (7) and (10). Equation (7) is first evaluated using only the first term of (10), i.e., using only the expected waveform, to give J_{eff1} . Equation (7) is evaluated again using both terms in (10), i.e., using both expected and variance waveforms, to give J_{eff2} . The percentage contribution of the variance waveform to J_{eff} is then measured as $\frac{J_{eff2} - J_{eff1}}{J_{eff1}} \times 100$. The results are tabulated in the fifth column of Table 1 for a number of test examples. The results clearly establish the importance of the variance waveform in addition to the expected waveform.

Finally, we illustrate the speed performance of CREST with the variance estimation built in. Table 1 shows also the speed comparisons between CREST and SPICE for all the examples presented above. The speedup becomes much better for larger circuits (1529X for

the adder and 11595X for the multiplier). In fact, the speedup should grow exponentially, because an exponential number of deterministic simulation runs are replaced by a single probabilistic simulation run. We point out the case of the multiplier circuit (the largest circuit in the table) with the heuristic CREST run (last row in Table 1); considering the excellent waveform comparison in Fig. 9 along with the dramatic speedup of 11595X in Table 1, this establishes the feasibility of solving large VLSI chips.

Table 1. Execution time comparisons. Time is in CPU seconds on a VAX-11/780; size refers to the number of transistors.

Circuit	Size	SPICE	CREST	Variance Contribution	Speedup
Nand	4	41.75	0.90	110%	46X
Complex	6	244.15	1.09	185%	224X
Xor	16	456.00	3.13	236%	146X
Adder	54	32620.42	21.33	107%	1529X
Multiplier	648	697530.88*	1871.99	219%	373X
Multiplier	648	697530.88*	60.16†	200%	11595X

*Estimated (2^8 times the cost of a typical logical SPICE run).

†Heuristic CREST run, all others are full accuracy CREST.

6. Summary and Conclusions

We have discussed the problem of estimating the median time-to-failure (MTF) due to electromigration in the power and ground busses of CMOS VLSI circuits. In this task we have verified that including the *variance waveform* of the current, in addition to the *expected waveform* derived in [1], further improves the accuracy of MTF estimation. This was done by showing that the variance contribution to the MTF estimate can be in the range of 100% to 200% relative to that of the expected current waveform. We have described a novel technique for deriving the variance waveform, and its implementation in the probabilistic simulator CREST. The results of several CREST runs have been presented, and they show good waveform agreement with SPICE, as well as excellent speedups over traditional approaches - a speedup of over 11000X was demonstrated on a 648-transistor circuit.

This work proves that the expected and variance waveforms of the stochastic current model are : (1) essential to derive an accurate MTF value, and (2) can be efficiently derived

using the probabilistic simulation approach. The main advantage of this approach is the ability to handle large circuits by replacing an exponential number of deterministic simulation runs with a single probabilistic simulation run. Without such a technique, analyzing the reliability of large circuits would seem to be an impossibility.

References

- [1] F. Najm, R. Burch, P. Yang, and I. Hajj, "CREST - a current estimator for CMOS circuits," *IEEE International Conference on Computer-Aided Design*, Santa Clara, CA, pp. 204-207, November 7-10, 1988.
- [2] J. D. Venables and R. G. Lye, "A statistical model for electromigration induced failure in thin film conductors," *Proc. 10th Annual IEEE Reliability Physics Symposium*, Las Vegas, NV, pp. 159-164, April 5-7, 1972.
- [3] J. W. McPherson and P. B. Ghate, "A methodology for the calculation of continuous DC electromigration equivalents from transient current waveforms," *The Electrochemical Society, Proc. Symp. on Electromigration of Metals*, New Orleans, LA, pp. 64-74, Oct. 7-12, 1984.
- [4] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd Edition. New York, NY: McGraw-Hill Book Co., 1984.
- [5] F. Najm, "Probabilistic simulation for reliability analysis of VLSI circuits," Ph.D. thesis, Dept. of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, June 1989.
- [6] R. Burch, F. Najm, P. Yang, and D. Hocevar, "Pattern-independent current estimation for reliability analysis of CMOS circuits," *25th ACM/IEEE Design Automation Conference*, Anaheim, CA, pp. 294-299, June 12-15, 1988.

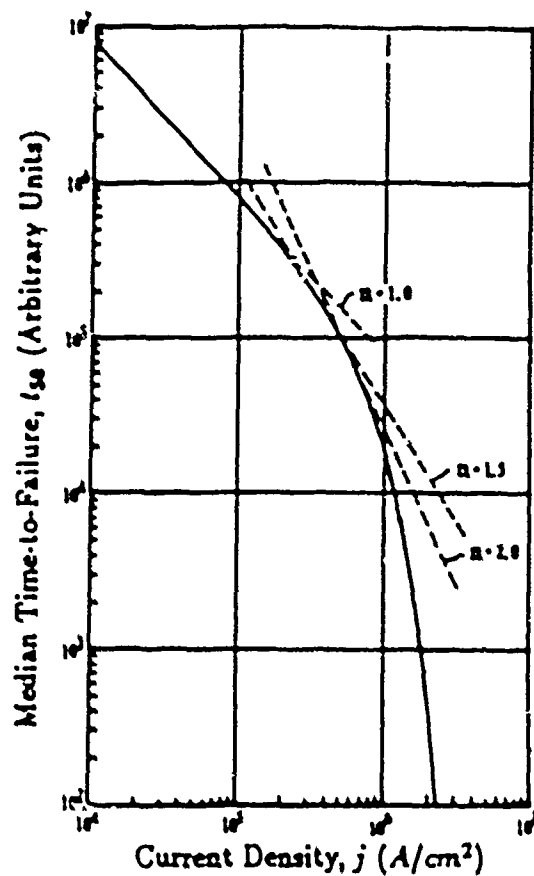


Figure 1: The dependence of MTF on current density, reproduced for convenience from [2]. The dashed lines show the results of the approximation $t_{50} \propto j^{-n}$ for $n = 1, 3/2$, and 2 .

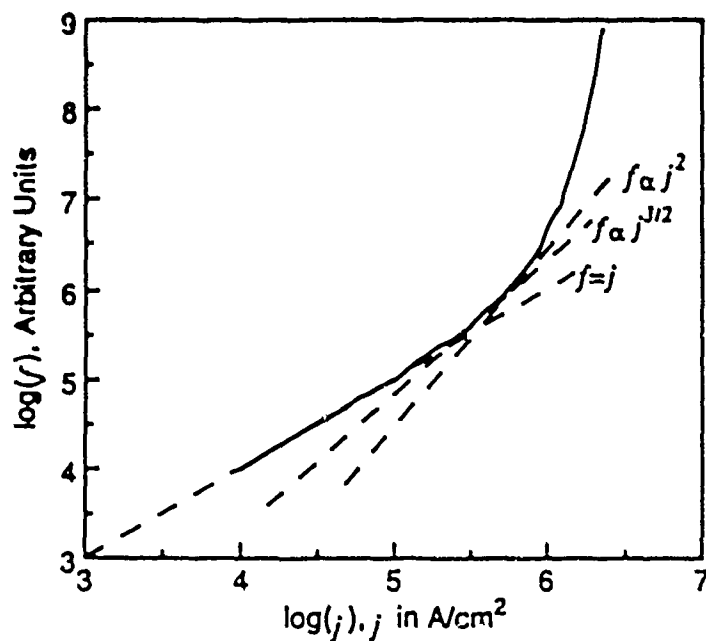


Figure 2: A plot of $f(j)$, obtained from Fig. 1 by inverting and appropriately scaling the ordinate axis.

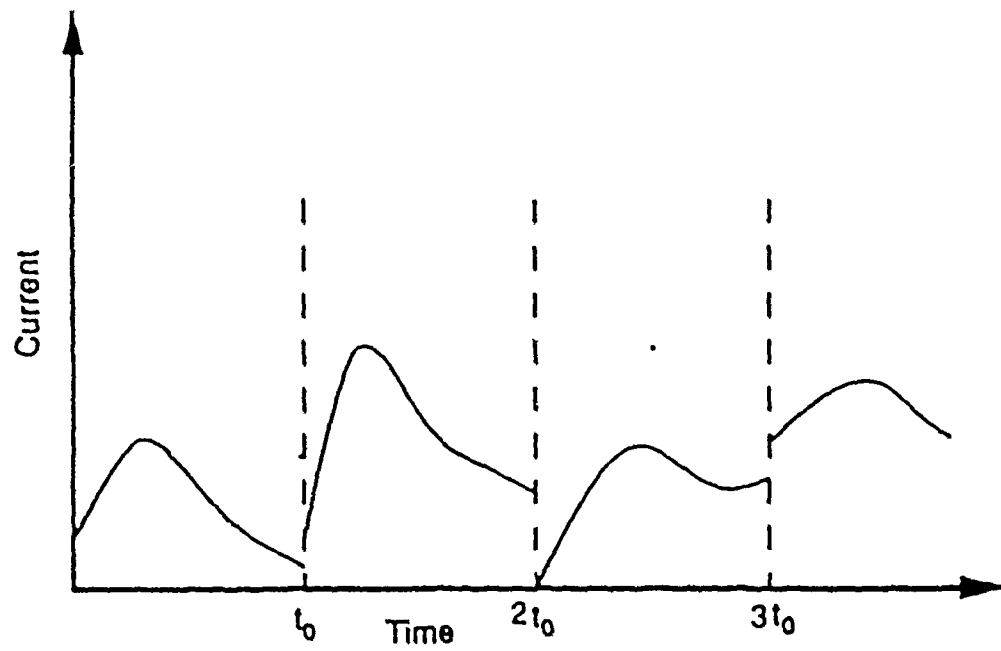


Figure 3: A (non-stochastic) current waveform, $j(t)$, built as a sequence of the waveforms $j_k(t)$, each occurring with its assigned probability P_k .

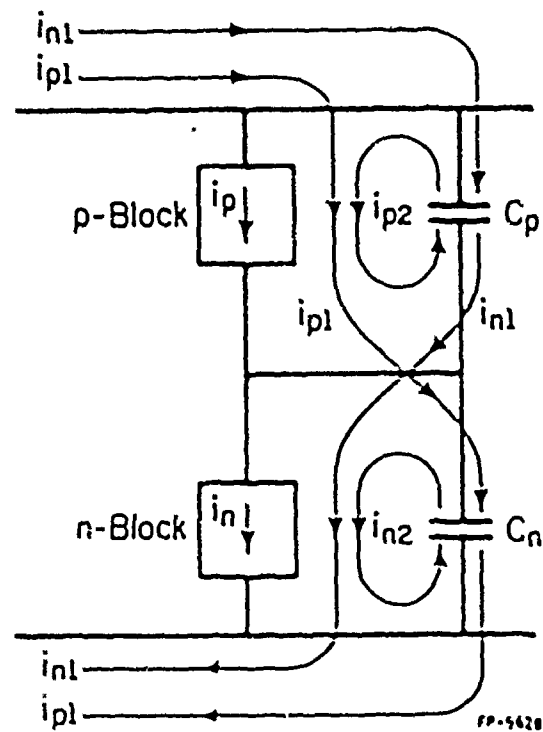


Figure 4: A generic CMOS gate structure.

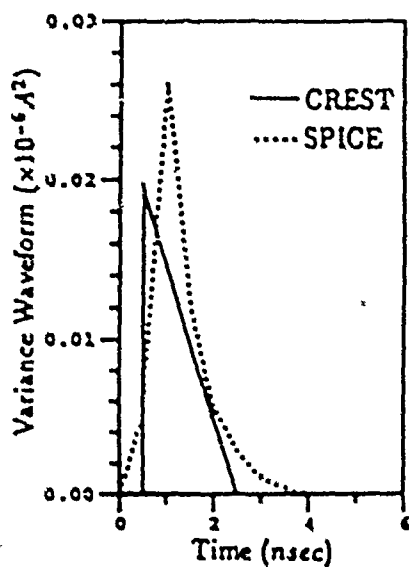


Figure 5: CREST variance pulse result for a 2-input CMOS nand gate, compared to SPICE.

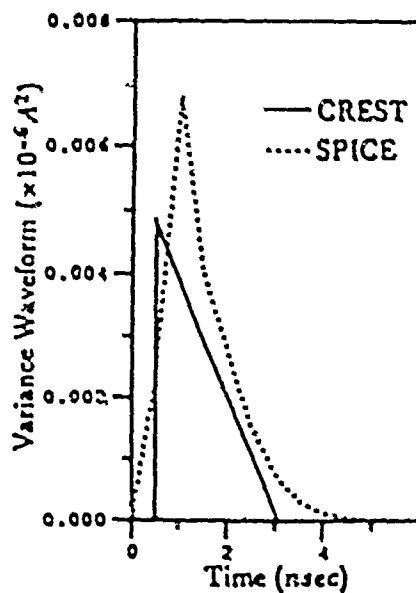


Figure 6: Variance results for a 3-input complex CMOS gate.

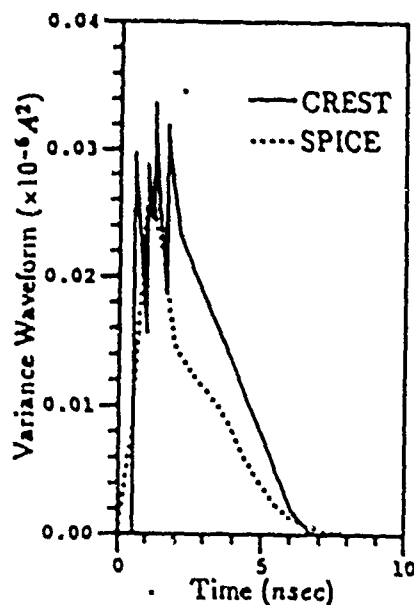


Figure 7: Variance results for a 16-MOSFET exclusive-or (xor) CMOS circuit.

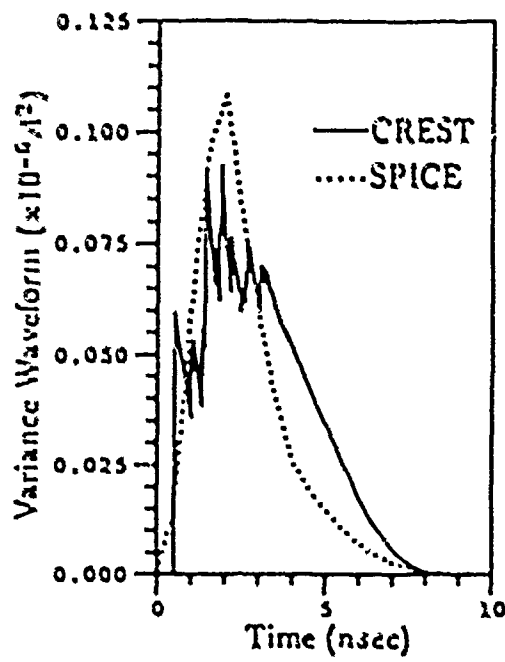


Figure 8: Variance results for a 54-MOSFET 2-bit ripple adder CMOS circuit.

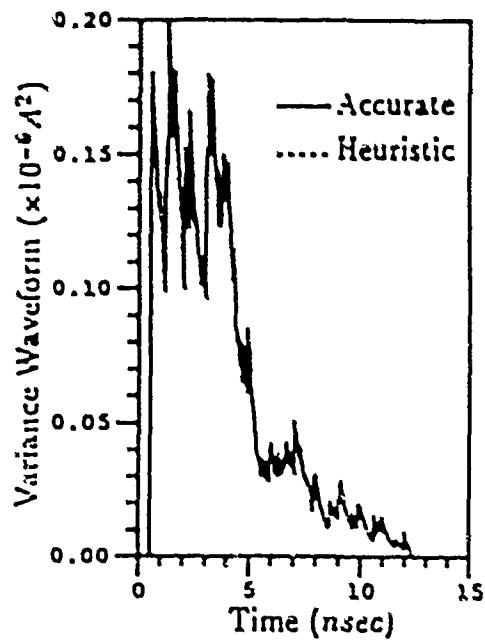


Figure 9: Variance results for a 648-MOSFET 4-bit parallel multiplier CMOS circuit.