Statistical Estimation of Circuit Timing Vulnerability Due to Leakage-Induced Power Grid Voltage Drop[*]

Imad A. Ferzli
Department of ECE
University of Toronto
Toronto, Ontario, Canada
ferzli@eecg.utoronto.ca

Farid N. Najm
Department of ECE
University of Toronto
Toronto, Ontario, Canada
f.najm@utoronto.ca

ABSTRACT

Statistical Vt variations lead to large variations of leakage current, which cause statistical voltage drops on the power grid that can affect circuit timing. We propose a statistical analysis technique whereby variances of the leakage currents are used to estimate the susceptibility to timing violations due to leakage-induced voltage drops.

1. INTRODUCTION

Managing leakage is among the most eminent challenges to be hurdled by the integrated circuits industry, with the advent of deep-submicron technologies in the nanometer regime. High leakage comes as a price tag on better performance and reduced active power, as a result of the need to scale the MOSFET threshold voltage, (V_{th}), to accompany the reduction in supply voltages (V_{dd}) and oxide thickness. This is because the decrease of V_{th} translates as an exponential increase in subthreshold leakage current, I_{off}, reportedly as high as about 5X per generation [6], with total leakage current in dense high-performance chips forecast to be about half the total chip current [13]. In today's 1.2 V, 0.13 μm technologies, V_{th} is about 0.3 V (25% of V_{dd}). Compare these figures WIth older 1μm technologies, using a supply voltage of 5V and having a threshold voltage of about 0.8 V (16% of V_{dd}) [26]. Clearly, the scaling trends of threshold voltage do not keep up with those of supply voltage.

The disparity in scaling the supply and threshold voltages is specially consequential to pursuing agressive designs, inducing tighter design margins, thus placing process variations under deep scrutiny. In particular, since the gap between V_{dd} and V_{th} has narrowed, variations in the level of supply voltage become very significant when it comes to meeting timing closure. The network that delivers supply voltage to circuit devices is referred to as the *power grid*. The consequences of power grid voltage level variations in the estimation of a circuit timing vulnerability are being currently researched [2, 21], and it has been reported in an industrial survey that 50% of chips would fail to meet their timing requirements if no power grid verification were done prior to tapeout [1]. However, modern power grid verification tools fail to account for the contribution of leakage to the voltage drop on the power grid. The present paper addresses this issue.

If leakage were uniform or predictable across the chip, then the contribution of leakage to the supply voltage drop can be simply accounted for by adding a deterministic leakage-induced component to the total voltage drop on the power grid. However, this is not the case and process variations in V_{th} and the transistor channel length cause exponentially larger variations in leakage currents, as a direct consequence of the exponential dependence of leakage currents on V_{th} or channel length [19, 24]. In 0.1 μm technology, it is possible to get 30 mV standard deviation in V_{th} [15]. For whole chips, leakage variations have been *measured* at almost 20X [14]. These variations are also known to have a significant within-die, local component [4, 5, 19], so that transistors in close proximity on the layout can have significant variations in their leakage currents. This is also referred to as mismatch and its effect on delay has been studied [8].

We will assume statistical variability of leakage currents, *i.e.*, that the statistics of leakage currents are available, with the knowledge that both Monte Carlo and analytical techniques can be used to obtain this information [24]. While worst-case analysis is necessary when considering power grid voltage drop due to global, correlated die-to-die variations [4] in leakage currents, intra-die (or within-die) variations require statistical analysis in order to avoid overly pessimistic conclusions. There are no good tools today to estimate this statistical voltage drop on the power grid due to leakage current variations.

In this work, we consider leakage currents to be random and proceed to analyze the grid on that basis. The focus is placed on within-die variations, which have been receiving a great deal of attention due to the impact they are forecast to have in future technologies. In [10], a numerical Monte Carlo approach was proposed to calculate partially the second order statistics of the power grid node voltages in response to within-die leakage current variations. This approach was extended in [11] to formulate a statistical verification methodology for the power grid considering leakage current variations, on an individual node basis. This methodology, however, did not consider joint statistics of the voltage drops on the grid nodes, which would be needed when simultaneous voltage drop conditions (joint conditions on a set of nodes) are to be checked against. Joint nodal verification is essential when checking the supply voltage levels simultaneously on a critical path. This paper proposes a way to deal with this situation. The result of this work is to extract joint statistics of the leakage-induced component

of the supply voltage drop on a critical path that indicates the vulnerability of a circuit to timing failure due to leakage currents. These statistics may then be fed to a tool such as the one presented in [2] when estimating the susceptibility of a circuit to timing violations.

2. PROBLEM STATEMENT

Designers may be able to determine a configuration of supply voltage variations that creates a worst-case effect on the timing performance of the circuit [2] and in practice set guidelines on the maximum allowable voltage drop levels on the power grid (*e.g.*, 10% of V_{dd}). From the perspective of IR drop on the grid (due to the grid's resistive nature), grid verification tools today account for switching current effects on the grid, ignoring the leakage component.

The power grid has the property of being a linear system (see section 3.1), therefore the grid response, *i.e.*, its voltage drops due to circuit currents, can be obtained simply by adding up the grid response to individual current components. In particular, both active and leakage current effects ought to be integrated in a comprehensive grid verification tool, and the grid response to the total current can be obtained by superposition of individual responses to both switching and leakage currents (static and dynamic).

Process-induced variations on the leakage component of the total current are relevant in the context of this work. It is standard practice to break down these process variations into *die-to-die* and *within-die* components [20, 7]. For the variations on a given parameter, the die-to-die component takes the same value for all instances of that parameter within a single chip, but differentiates among distinct chips. The within-die component causes variations within a given chip, depending on location.

We therefore model the total current as follows:

$$I = I_{\text{active}} + I_{\text{off,nom}} + I_{\text{off,dd}} + I_{\text{off,wd}}, \qquad (1)$$

where I_{active} refers to the active current, $I_{\text{off,nom}}$ is the *nominal* value of leakage current at the design point, $I_{\text{off,dd}}$ is the die-to-die component of leakage variations, and $I_{\text{off,wd}}$ is the within-die component of these variations. By linearity of the power grid, we can add the contributions of the four current components in (1) to get the total voltage drop on the power grid, as follows:

$$V = V_{I_{\text{active}}} + V_{I_{\text{off}},\text{nom}} + V_{I_{\text{off}},\text{dd}} + V_{I_{\text{off}},\text{wd}}. \qquad (2)$$

In general, die-to-die variations can be dealt with efficiently using case file analysis. This is also applicable for examining the response of the power grid due to die-to-die leakage current variations, since the grid is known to be *monotone* [17], *i.e.*, the voltage drops on all nodes increase by increasing the loading currents. When the current variations are correlated, this means that all currents may increase simultaneously, so that case file analysis may be performed reflecting realistic situations. This can be done at the (mean±3σ) corners of die-to-die variations, or at more agressive points (*e.g.*, the 50$^{\text{th}}$ percentile). Therefore, the analysis of random die-to-die variations can proceed by analyzing a set of deterministic cases.

The considerations are different for within-die variations. Due to their locality on the die area, the occurrence of worst-case deviations uniformly across the die is highly unlikely, and enumerating cases becomes prohibitive. Instead, statistical analysis needs to be performed.
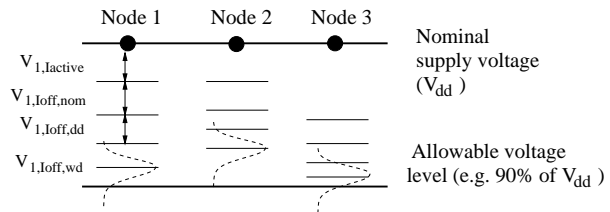

Figure 1: Grid verification on a critical path containing 3 nodes.

Fig. 1 illustrates the situation. Suppose, in this figure, that nodes 1, 2, and 3 of the power grid provide supply voltage connections to circuit elements lying on a critical path (referred to thereafter as *critical nodes*), and that for timing purposes, it was determined that the voltage levels on these nodes need to be simultaneously greater than 90% of V_{dd}. The figure illustrates the voltage drops contributed by active current, nominal leakage, die-to-die, and within-die variations. Within-die variations are shown to have a certain probability distribution, which necessarily precludes the ability to determine with 100% confidence whether the required voltage levels are greater than 90% of V_{dd}. The need for joint node voltage statistics is clear as we seek statistics on the simultaneous distributions of the voltage drop at critical nodes, in order to have the ability to explore the margin of within-die variations efficiently. The sequel focuses on the joint statistics of within-die variations of voltage drops.

3. JOINT STATISTICS OF POWER GRID NODE VOLTAGE DROPS

It is helpful to distinguish between two types of leakage in integrated circuits. A circuit certainly draws leakage current when it is in standby or sleep mode, what may be referred to as the *standby leakage*. The circuit also draws leakage current when it is active. Indeed, a logic gate draws leakage current any time that its supply is "on." Even inside a switching window, part of the current drawn from the supply may be attributed to leakage. The leakage drawn by the circuit during its active (non-standby) states, may be referred to as the *dynamic leakage*. The grid response to standby leakage may be obtained by a DC analysis of the grid, using only a resistive model, whereas response to dynamic leakage requires a transient analysis, using an RC or RLC model of the grid. Our statistical formulation to follow, directly applies to standby leakage. Dynamic leakage may be incorporated through calculation of the mean voltage drops on a transient basis, as discussed in section 4.1.

3.1 System Equations

We consider an RC model of the power grid, where each branch of the grid is represented by a resistor and where there exists a capacitor from every grid node to ground. In addition, some nodes have ideal current sources (to ground) representing the current drawn by the circuit tied to the grid at that point, and some nodes have ideal voltage sources (to ground) representing the connections to the external supply voltage. Let the power grid consist of $N + p$ nodes, where nodes $1, 2, \ldots, N$ have no voltage sources attached, and nodes $(N+1), (N+2), \ldots, (N+p)$ are the nodes with the voltage sources. Let c_k be the capacitance from every node k to ground. Let $i_k(t)$ be the current source connected

to node k, where the direction of positive current is from the node to ground. We assume that $i_k(t) \geq 0$ and that $i_k(t)$ is defined for every node $k = 1, \ldots, N$ so that nodes with no current source attached have $i_k(t) = 0, \forall t$. Let $\mathbf{i}(t)$ be the vector of all $i_k(t)$ sources, $u_k(t)$ be the voltage at node k, and $\mathbf{u}(t)$ be the vector of all $u_k(t)$ signals. Applying Modified Nodal Analysis (MNA) [23] leads to:

$$\mathbf{G}\mathbf{u}(t) + \mathbf{C}\dot{\mathbf{u}}(t) = -\mathbf{i}(t) + \mathbf{G}\mathbf{V_{dd}} \quad (3)$$

where $\mathbf{G}$ is an $N \times N$ conductance matrix, $\mathbf{C}$ is an $N \times N$ diagonal matrix of node capacitances, and $\mathbf{V_{dd}}$ is a constant vector each entry of which is equal to V_{dd}. Let $v_k(t) = V_{dd} - u_k(t)$ be the voltage drop at node k, and let $\mathbf{v}(t)$ be the vector of voltage drops, then (3) can be written as:

$$\mathbf{G}\mathbf{v}(t) + \mathbf{C}\dot{\mathbf{v}}(t) = \mathbf{i}(t) \quad (4)$$

This is a *revised* system equation which one can solve directly for the voltage drop values. Notice that the circuit described by (4) consists of the original power grid, but with all the voltage sources set to zero and all the current source directions reversed. In the following, we will mainly be concerned with this *modified power grid* and the revised system of equations (4). In cases when the circuit is in a standby state, where all the currents are constant, the circuit response is obtained using a DC analysis. The DC equivalent of (4) is readily seen as:

$$\mathbf{G}\mathbf{V} = \mathbf{I} \quad (5)$$

3.2 Joint Probability Distribution of Node Voltages

DC analysis of the power grid leads to an analytical characterization of the probability distribution function of the node voltage drops under within-die leakage current variations. DC analysis is automatically true for the case of standby leakage. Thus, our analysis is exact for standby leakage. For dynamic leakage, since the leakage current of a logic gate is constant when it is not switching, then DC analysis may be acceptable in practice, especially since we will include *some* dynamics of the system through the computation of the mean response (section 4.1 below).

The DC representation of the power grid system given in (5) can be written as follows:

$$\mathbf{V} = \mathbf{G}^{-1}\mathbf{I}, \quad (6)$$

where we consider $\mathbf{I}$ to be the *random vector* representing within-die leakage variations. Correspondingly, $\mathbf{V}$ is the vector of grid voltage drops due to $\mathbf{I}$. From (6), it is clear that the voltage drop at an arbitrary node i can be expressed as:

$$V_i = q_{i1}I_1 + \cdots + q_{iN}I_N, \quad (7)$$

where q_{ij} is the (i,j)th entry of $\mathbf{G}^{-1}$. As (7) shows, the voltage drop at any node is a linear combination of the leakage currents loading the grid.

Our purpose is to characterize intra-die variations in leakage currents. These are due to intra-die variations in the transistor threshold voltages, which are in turn modeled as Gaussian [19], implying *lognormal* [28] distributions for leakage currents [24], by virtue of the exponential relation between the transistor leakage current and its threshold voltage, which can be seen from the following BSIM 3 device model of leakage current, in terms of physical and electrical

parameters [24]:

$$I_{\text{off}} = I_0 \cdot \exp\left((V_{gs} - V_{th})/nV_T\right)\left(1 - \exp\left(-V_{ds}/V_T\right)\right), \quad (8)$$

where V_T is the thermal voltage and I_0 is a constant for our purposes.

If the I_j are lognormal, then the $q_{ij}I_j$ are also lognormal. Hence, under the assumption of statistical independence of intra-die leakage variations, the node voltages (7) are each a summation of independent lognormal random variables (RVs). Sums of independent lognormal variables have been extensively studied and characterized in the literature pertaining to communications, and it was found that such sums can be accurately captured by another lognormal RV [3]. Therefore, we model the distribution of a node voltage drop as lognormal. For joint node voltage statistics, we *propose* a multivariate lognormal distribution.

Recall that a random variable X is said to be lognormal if $Y = \ln(X)$ is Gaussian. Let μ_Y and σ_Y be the mean and standard deviation of Y. Then the probability density function (pdf), $f_X(.)$, of X is given by [28]:

$$f_X(x) = \exp\left(-\frac{[\ln(x) - \mu_Y]^2}{2\sigma_Y^2}\right) / \left(\sqrt{2\pi}\sigma_Y x\right), \quad x > 0. \quad (9)$$

The multivariate lognormal generalizes the above univariate distribution. A random vector $\mathbf{X}$ is multivariate lognormal when its logarithm is jointly normal. Let $\mathbf{Y} = \ln(\mathbf{X})$, where $[\mathbf{Y}]_i = \ln([\mathbf{X}]_i)$, with $[.]_i$ denoting the ith component of a vector. If k is the dimension of $\mathbf{X}$, then for a k-vector $\mathbf{x} > 0$, the joint probability density function of $\mathbf{X}$ can be written as [16]:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\exp\left(-\frac{1}{2}\left(\ln(\mathbf{x}) - \mu_{\mathbf{Y}}\right)^T \Sigma_{\mathbf{Y}}^{-1} \left(\ln(\mathbf{x}) - \mu_{\mathbf{Y}}\right)\right)}{(2\pi)^{k/2} |\Sigma_{\mathbf{Y}}|^{1/2} [\mathbf{x}]_1 \cdots [\mathbf{x}]_k}, \quad (10)$$

where $\mu_{\mathbf{Y}}$ and $\Sigma_{\mathbf{Y}}$ refer to the mean vector and *covariance matrix* of the random vector $\mathbf{Y}$, $|.|$ refers to a matrix determinant, $(.)^T$ to the transpose operator, and $\mu_{\mathbf{Y}}$ to the mean vector of $\mathbf{Y}$. Denoting by $[.]_{ij}$ the $(i,j)^{th}$ entry in a matrix, and by cov$(.)$ the covariance between two random variables, $\Sigma_{\mathbf{Y}}$ is defined as [25]:

$$[\Sigma_{\mathbf{Y}}]_{ij} = \begin{cases} \text{cov}([\mathbf{Y}]_i, [\mathbf{Y}]_j) & , \quad i \neq j \\ \sigma_{[\mathbf{Y}]_i}^2 & , \quad i = j \end{cases}$$

A few observations are in order at this point. First note that the pdf of a multivariate lognormal is defined only when the corresponding normal vector has a non-singular covariance matrix. Second, because of its close association to the multivariate normal, full statistics of node voltage drops can be captured by second-order statistics (means and a covariance matrix). Finally, any k-vector taken from an $n-$dimensional lognormal vector ($k \leq n$) is (multivariate) lognormal itself.

Now we can put the above to use in our problem. Let $\mathbf{V_k}$ be the k-vector of critical nodes of the power grid. k is expected to be much smaller than N, the dimension of the grid. $\mathbf{V_k}$ will be modeled as a lognormal vector. Let $\mathbf{W_k} = \ln(\mathbf{V_k})$ and let $\mu_{\mathbf{W_k}}$ and $\Sigma_{\mathbf{W_k}}$ be respectively the mean vector and the covariance matrix of $\mathbf{W_k}$. If $\mathbf{v} = [v_1 \cdots v_k]^T > 0$,

then the pdf of $\mathbf{V_k}$ at $\mathbf{v}$ is given by:

$$f_{\mathbf{V_k}}(\mathbf{v}) = \frac{\exp\left(-\frac{1}{2}\left(\ln(\mathbf{v}) - \mu_{\mathbf{W_k}}\right)^T \Sigma_{\mathbf{W_k}}^{-1}\left(\ln(\mathbf{v}) - \mu_{\mathbf{W_k}}\right)\right)}{(2\pi)^{k/2}\left|\Sigma_{\mathbf{W_k}}\right|^{1/2} v_1 \cdots v_k}$$
(11)

Let $V_{iv} = V_{im} - V_{i,I_{\mathrm{active}}} - V_{i,I_{\mathrm{off}},\mathrm{nom}} - V_{i,I_{\mathrm{off}},\mathrm{dd}}$, where V_{im} is the maximum allowable voltage drop at node i and the other voltage levels on the right-hand side are as defined in (2) and Fig. 1. The probability of meeting the timing requirements becomes:

$$\mathcal{P}\{V_{i,I_{\mathrm{off}},\mathrm{wd}} \leq V_{iv}, \forall i = 1, \ldots, k\} =$$
$$\int\limits_{0}^{V_{kv}} \cdots \int\limits_{0}^{V_{1v}} f_{\mathbf{V_k}}(\mathbf{v})\mathrm{d}v_1 \cdots \mathrm{d}v_k$$
(12)

We recognize that the difficulties associated with performing the above operations are numerical, as multivariate manipulations are tedious. The first difficulty is to estimate the parameters of the multivariate lognormal distribtion, for which we propose a Monte Carlo approach (section 4.2). Since we expect the number k of critical nodes to be much smaller than the size N of the grid, and indeed only a few tens of nodes may need to be considered, the dimensions are manageable for storing non-sparse covariance matrices, as well as calculating the determinant and matrix inverse. The evaluation of the multidimensional integral given in (11) must be handled with numerical techniques, but is alleviated by the fact that the integrand is a smooth function and that the domain of integration is a "hyper-rectangle". In our experiments, we performed the multiple integration using Monte Carlo methods.

4. ESTIMATION OF THE DISTRIBUTION PARAMETERS

4.1 Estimation of The Mean Vector of Voltage Drops

Since the system (4) is linear, then due to linearity of the mean ($E[\cdot]$) operator [22], one can write:

$$\mathbf{G}E[\mathbf{v}(t)] + \mathbf{C}\frac{d}{dt}E[\mathbf{v}(t)] = E[\mathbf{i}(t)]$$
(13)

Thus, solving the system (4) once, using simply the means of the within-die leakage current variations as inputs, yields mean voltage drops at all the nodes obtained from the dynamic model of the grid.

4.2 Estimation of The Covariance Matrix of Voltage Drops

We start by deriving an analytical expression for the covariance of the voltage at every node. Under static conditions, the system (13) is simplified to its DC version:

$$\mathbf{G}E[\mathbf{V}] = E[\mathbf{I}]$$
(14)

We now combine (5) and (14) to yield:

$$\mathbf{G}(\mathbf{V} - E[\mathbf{V}]) = \mathbf{I} - E[\mathbf{I}]$$
(15)

Multiplying each side by its transpose and applying the ex-

pected value operator to each side, leads to:

$$\mathbf{G}E\left[(\mathbf{V} - E[\mathbf{V}])(\mathbf{V} - E[\mathbf{V}])^T\right]\mathbf{G}^T =$$
$$E\left[(\mathbf{I} - E[\mathbf{I}])(\mathbf{I} - E[\mathbf{I}])^T\right]$$
(16)

We recognize the expectations as being simply covariance matrices [22], so that the above result can be rewritten as:

$$\mathbf{G}\Sigma_{\mathbf{V}}\mathbf{G}^T = \Sigma_{\mathbf{I}}$$
(17)

Since $\mathbf{G}$ is symmetric, $\mathbf{G} = \mathbf{G}^T$. Therefore, (17) becomes:

$$\Sigma_{\mathbf{V}} = \mathbf{G}^{-1}\Sigma_{\mathbf{I}}\mathbf{G}^{-1}$$
(18)

Under the assumption of statistical independence of leakage currents, implying that $\Sigma_{\mathbf{I}}$ is diagonal, it can be seen from (18) that:

$$[\Sigma_{\mathbf{V}}]_{ij} = \left[\mathbf{G}^{-1}\right]_{i1}\left[\mathbf{G}^{-1}\right]_{j1}[\Sigma_{\mathbf{I}}]_{11} + \cdots +$$
$$\left[\mathbf{G}^{-1}\right]_{iN}\left[\mathbf{G}^{-1}\right]_{jN}[\Sigma_{\mathbf{I}}]_{NN}$$
$$= q_{i1}q_{j1}\sigma_{I_1}^2 + \cdots + q_{iN}q_{jN}\sigma_{I_N}^2,$$
(19)

where q_{ij} is as defined in (7) and $\sigma_{I_i}^2$ is the variance of the ith leakage current, that is, the ith diagonal entry of the covariance matrix of leakage currents.

The solution of (19) requires full knowledge of the inverse of $\mathbf{G}$, $i.e.$, the inverse of the covariance matrix of the entire power grid, even when only the sub-covariance matrix corresponding to critical nodes is required. To manage this problem, we extend the method presented in [10] to handle covariance matrices.

Observe that (19) is a weighted summation. Let $S = \sum_{i=1}^{N}\sigma_{I_i}^2$, $p_i = \sigma_{I_i}^2/S$, and $\sigma_{ij}^{(\mathbf{V_k})} = [\Sigma_{\mathbf{V}}]_{ij}$, when nodes i and j on the power grid are critical. Then (19) can be rewritten as:

$$\sigma_{ij}^{(\mathbf{V_k})} = S\sum_{l=1}^{N} p_l(q_{il}q_{jl}).$$
(20)

Since $\sum_{k=l}^{N} p_l = 1$ and $p_l \geq 0$, $l = 1\cdots N$, then we can view the p_l weights as being $probability\ values$ associated with the $q_{il}q_{jl}$ values, so that the summation in (20) becomes the mean (weighted average) of all the $q_{il}q_{jl}$ values in the ith and j^{th} rows. If we define an RV $\mathbf{q_{ij}}$ as being a discrete RV that takes the values $q_{ijl} = q_{il}q_{jl}$ with probabilities p_l, $l = 1, 2, \ldots, N$, then we can write (20) as:

$$\sigma_{ij}^{(\mathbf{V_k})} = SE[\mathbf{q_{ij}}]$$
(21)

Let the mean of $\mathbf{q_{ij}}$ be $\mu_{ij} = E[\mathbf{q_{ij}}]$ and its variance be v_{ij}. We can now use methods of $mean\ estimation$ from statistics, basically Monte Carlo random sampling [27], in order to estimate the $population\ mean\ \mu_{ij}$ using the mean of a much smaller sample (say, of size $n \ll N$) from the population, $i.e.$, using the $sample\ mean$.

Using a weighted random number generator, we generate according to the probabilities p_l a sequence of indices of columns of $\mathbf{G}^{-1}$ to be included in the sample. From these, we form the following sample mean for every pair of rows i and j for which nodes i and j are critical:

$$\bar{q}_{ij} = \frac{1}{n}\sum_{l \in \mathcal{L}} q_{ijl}$$
(22)

where $\mathcal{L}$ is the set of indices included in the random sample. Note that with a total of k nodes, there will be a total of $\binom{k}{2} = k(k-1)/2$ pairwise covariances (off-diagonal entries) and k variances, for a total of $k(k+1)/2$ distinct entries to be estimated. By symmetry of the covariance matrix, the remaining $k(k-1)/2$ entries can be deduced, but do not need to be stored.

We also compute the *sample standard deviation* of $\mathbf{q_{ij}}$, $s_{ij} \geq 0$ given by:

$$s_{ij}^2 = \frac{1}{n-1} \sum_{l \in \mathcal{L}} (q_{ijl} - \bar{q}_{ij})^2 = \frac{n\left(\sum_{l \in \mathcal{L}} q_{ijl}^2\right) - \left(\sum_{l \in \mathcal{L}} q_{ijl}\right)^2}{n(n-1)} \tag{23}$$

Note that s_{ij}^2 is an estimator of v_{ij}. Now $\bar{q}_{ij}$ itself, being a sample mean, can be considered as an RV, with mean μ_{ij} (since the sample mean is an unbiased estimator of the mean of a random variable [28]) and variance s_{ij}^2/n (for large n). Then, we have that:

$$\hat{\sigma}_{ij}^{(\mathbf{V_k})} = S\bar{q}_{ij} \tag{24}$$

is an unbiased estimator of $\sigma_{ij}^{(\mathbf{V_k})}$ (*i.e.* $E[\hat{\sigma}_{ij}^{(\mathbf{V_k})}] = SE[\bar{q}_{ij}] = \sigma_{ij}^{(\mathbf{V_k})}$), with variance $S^2 s_{ij}^2/n$. Furthermore, by the *central limit theorem* [28], $\bar{q}_{ij}$ will be approximately normally distributed, so that the RV:

$$\frac{\sigma_{ij}^{(\mathbf{V_k})} - \hat{\sigma}_{ij}^{(\mathbf{V_k})}}{\frac{S s_{ij}}{\sqrt{n}}}$$

is normal with 0 mean and unit variance [27], for large n. Fig. 2 illustrates the sampling process for estimating the covariance between nodes 1 and 3. The convergence of the sample mean is directly related to the sample standard deviation.

We can use tables of the standard normal to establish how large n should be in order for the sample standard deviation of $\bar{q}_{ij}$ to be small enough for it to be a viable estimator of μ_{ij}, with a certain confidence, and up to a predefined tolerance [27]. For example, if it is desired to have $(1-\alpha) \times 100\%$ confidence (where α is a small positive number, $0 < \alpha < 1$) that the following is true:

$$|\bar{q}_{ij} - \mu_{ij}| < \epsilon, \tag{25}$$

then it is known from sampling theory [27] that n should be larger than n_0 where:

$$n_0 = \left(\frac{z_{\alpha/2} s_{ij}}{\epsilon}\right)^2 \tag{26}$$

where $z_{\alpha/2}$ is such that the area to the right of it under the pdf of the standard normal curve is equal to $\alpha/2$. Thus, for instance, for 90% confidence, $\alpha = 0.05$ and $z_{\alpha/2} = 1.96$; for 99% confidence, $\alpha = 0.01$ and $z_{\alpha/2} = 2.575$. In practice, one samples until n is larger than 30 or 50 or so, then starts to use (26) to monitor convergence. Observe that tables are not needed to compute $z_{\alpha/2}$, a task that can be done by software (*e.g.*, using the erf() function) and that $z_{\alpha/2}$ needs to be evaluated only once.

Next, we need to choose a meaningful bound ϵ for use in (26). We opted to relate the value of the error in the estimation of the covariance matrix entries to the reference voltage V_{dd}, as we did for the variance estimation in [10].
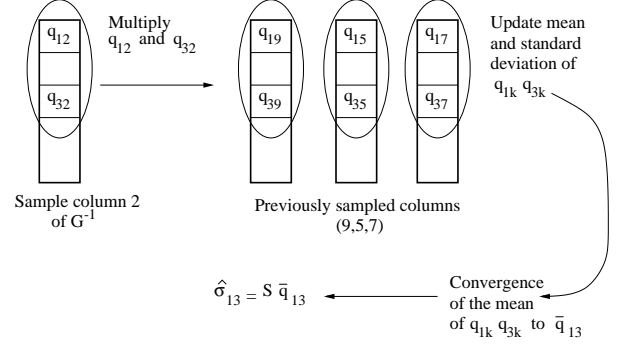

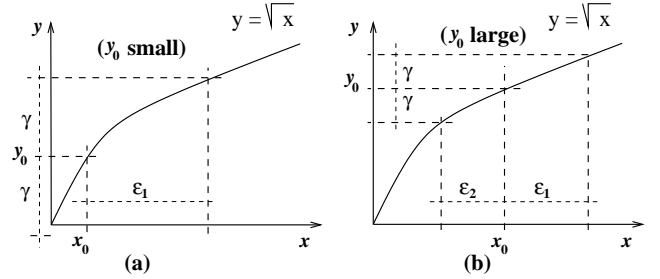Figure 2: Estimating the covariance between nodes 1 and 3 on the power grid.

Figure 3: Choice of error bound ϵ in the two cases when (a) y_0 is small and (b) y_0 is not so small.

Let δ be a small positive number, $0 < \delta < 1$. We define our error bound as follows:

$$\left| \sqrt{\sigma_{ij}^{(\mathbf{V_k})}} - \sqrt{S\bar{q}_{ij}} \right| \leq \delta V_{dd} \tag{27}$$

In other words, we want to find ϵ, to be used in (26), as a function of δ in order for the following to be true:

$$|\bar{q}_{ij} - \mu_{ij}| < \epsilon \implies \left| \sqrt{S\bar{q}_{ij}} - \sqrt{S\mu_{ij}} \right| < \delta V_{dd} \tag{28}$$

To simplify the notation, let $x = \mu_{ij}$ and $x_0 = \bar{q}_{ij}$. Also, let $y = \sqrt{\mu_{ij}} = \sqrt{x}$ and $y_0 = \sqrt{\bar{q}_{ij}} = \sqrt{x_0}$, and let $\gamma = \delta V_{dd}/\sqrt{S}$. Notice that $\gamma > 0$. We want to find ϵ in terms of δ so that:

$$|x - x_0| < \epsilon \implies |y - y_0| < \gamma \tag{29}$$

There are two cases to consider, according to whether y_0 is small or not, as shown in Fig. 3. When y_0 is small, small enough so that $y_0 - \gamma < 0$, then (since $y > 0$ in all cases) in order to guarantee that $|y - y_0| < \gamma$, it is sufficient to impose an upper bound on y in the simple form $y - y_0 < \gamma$. This is achieved by imposing an upper bound on x in the simple form $x - x_0 < \epsilon_1$, where ϵ_1 is the corresponding upper bound on $(x - x_0)$, as shown in Fig. 3(a). Let $\Delta x = x - x_0$ and $\Delta y = y - y_0$. Since $y^2 = x$, then $\Delta(y^2) = \Delta x$, and since $\Delta(y^2) = (y_0 + \Delta y)^2 - y_0^2 = (\Delta y)^2 + 2y_0 \Delta y$, then $\Delta x = ((\Delta y)^2 + 2y_0 \Delta y)$. For $\Delta y = \gamma$, $\epsilon_1 = \Delta x = (\gamma^2 + 2y_0\gamma) > 0$. In order to guarantee (29), in this case, we need to set $\epsilon = \epsilon_1 = (2y_0\gamma + \gamma^2)$.

When y_0 is not so small, i.e., when $y_0 - \gamma > 0$, then we need to consider both upper bounds and lower bounds, as shown in Fig. 3(b). If ϵ_2 is the lower bound on $(x - x_0)$, as shown in the figure, then we may compute ϵ_2 by setting $\Delta y = -\gamma$, which leads to $\Delta x = (\gamma^2 - 2y_0\gamma) < 0$, and we set $\epsilon_2 = -\Delta x = (2y_0\gamma - \gamma^2) > 0$. Since $\epsilon_2 < \epsilon_1$, as is obvious from the expressions found for each, then, in order

to guarantee (29) in this case, we need to set $\epsilon = \epsilon_2 = (2y_0\gamma - \gamma^2)$.

Notice that the condition $y_0 > \gamma$ translates to $\bar{q}_{ij} > \delta^2 V_{dd}^2/S$. In summary, then:

$$\epsilon = \begin{cases} \frac{\delta V_{dd}}{S}\left(2\sqrt{S\bar{q}_{ij}} + \delta V_{dd}\right) & \text{if } \bar{q}_{ij} < \delta^2 V_{dd}^2/S, \\ \frac{\delta V_{dd}}{S}\left(2\sqrt{S\bar{q}_{ij}} - \delta V_{dd}\right) & \text{if } \bar{q}_{ij} > \delta^2 V_{dd}^2/S. \end{cases} \quad (30)$$

Notice also that, in either case, $\epsilon > \delta^2 V_{dd}^2/S$. Plugging (30) into (26) leads to:

$$n_0 = \left(\frac{z_{\alpha/2}s_{ij}S/\delta V_{dd}}{\delta V_{dd} \pm 2\sqrt{S\bar{q}_{ij}}}\right)^2 \quad (31)$$

where the $+$ or $-$ sign depends on whether $\bar{q}_{ij}$ is smaller or larger than $\delta^2 V_{dd}^2/S$, respectively. To summarize, if it is desired to estimate $\sqrt{\sigma_{ij}^{(\mathbf{V_k})}}$ to within $\pm\delta V_{dd}$, with $(1-\alpha)\times 100\%$ confidence, then n must be larger than the threshold given by (31). This provides a useful trade-off between accuracy and speed, as more samples would be required for smaller δ.

4.3 Estimation of The Parameters of The Lognormal Distribution

The previous two subsections dealt with estimating the mean and covariance matrix of the vector of voltage drops due to leakage variations. In the terminology of subsection 3.2, these refer to the mean, $\mu_{\mathbf{V_k}}$, and covariance matrix, $\Sigma_{\mathbf{V_k}}$, of $\mathbf{V_k}$, whereas the parameters to be used in the expression of the probability distribution function given in (11) are $\mu_{\mathbf{W_k}}$ and $\Sigma_{\mathbf{W_k}}$, the mean vector and covariance matrix of the associated k-dimensional normal vector. We show how to obtain these parameters for the lognormal.

Suppose V_i and V_j $(i, j < k)$ are random voltage drops, taken from a joint k-dimensional lognormal distribution with parameters $\mu_{\mathbf{W_k}} = [\mu_1 \cdots \mu_k]^T$ and $[\Sigma_{\mathbf{W_k}}]_{ij} = \sigma_{ij}$. Here, we define $\sigma_{ij}^{(\mathbf{V_k})} = [\Sigma_{\mathbf{V_k}}]_{ij}$. Observe that at this point, we have estimated $\mu_{\mathbf{V_k}}$ and $\Sigma_{\mathbf{V_k}}$, but the $\mathbf{W_k}$ parameters are yet to be determined.

The following holds for the multivariate lognormal distribution [16]:

$$E[V_i] = e^{\mu_i + \frac{1}{2}\sigma_{ii}}, \quad (32)$$

$$E[V_i^2] = e^{2\mu_i + 2\sigma_{ii}}, \quad (33)$$

so that:

$$\begin{aligned} \sigma_{ii}^{(\mathbf{V_k})} &= E[V_i^2] - (E[V_i])^2 \\ &= e^{2\mu_i + 2\sigma_{ii}} - \left(e^{\mu_i + \frac{1}{2}\sigma_{ii}}\right)^2 \\ &= e^{2\mu_i + \sigma_{ii}}\left(e^{\sigma_{ii}} - 1\right) \\ &= (E[V_i])^2\left(e^{\sigma_{ii}} - 1\right). \end{aligned} \quad (34)$$

The only unknowns in (34) are the diagonal entries σ_{ii} of $\Sigma_{\mathbf{W_k}}$, which can thus be solved one by one. Then we plug these solved values into (32), and solve for the entries of $\mu_{\mathbf{W_k}}$.

We also have for the multi-lognormal distribution:

$$\sigma_{ij}^{(\mathbf{V_k})} = e^{\mu_i + \mu_j + \frac{1}{2}(\sigma_{ii} + \sigma_{jj})}(e^{\sigma_{ij}} - 1). \quad (35)$$

The only unknowns left at this stage are the σ_{ij} $(i \neq j)$, and these can again be found from (35). Hence, we have

12
10
8
6
4
2
0

Approximate values of the entries of the covariance matrix (normalized as % Vdd)

Exact values of the entries of the covariance matrix (normalized as % Vdd)
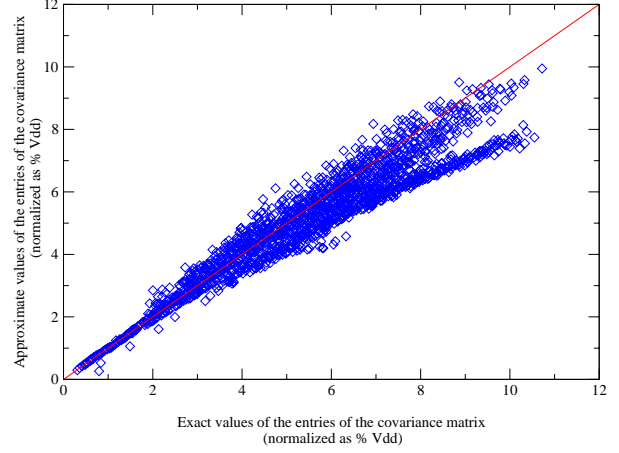
Figure 4: Correlation plot of the estimated versus exact entries of a full covariance matrix for a grid containing 300 nodes.

obtained the entries of $\mu_{\mathbf{W_k}}$ and $\Sigma_{\mathbf{W_k}}$ from the mean and covariance matrix of the node voltage drops.

The following sums up the procedure for obtaining $\mu_{\mathbf{W_k}}$ and $\Sigma_{\mathbf{W_k}}$ given $\mu_{\mathbf{V_k}}$ and $\Sigma_{\mathbf{V_k}}$:

```
for i = 1 : k
```
$$[\Sigma_{\mathbf{W_k}}]_{ii} = \ln\left(\frac{[\Sigma_{\mathbf{V_k}}]_{ii}}{[\mu_{\mathbf{V_k}}]_i^2} + 1\right)$$
$$[\mu_{\mathbf{W_k}}]_i = \ln\left([\mu_{\mathbf{V_k}}]_i\right) - \frac{1}{2}[\Sigma_{\mathbf{W_k}}]_{ii}$$
```
end for
```

```
for all i, j ≤ k, i ≠ j
```
$$[\Sigma_{\mathbf{W_k}}]_{ij} = \ln\left(\frac{[\Sigma_{\mathbf{V_k}}]_{ij}}{e^{[\mu_{\mathbf{W_k}}]_i + [\mu_{\mathbf{W_k}}]_j + \frac{1}{2}\left([\Sigma_{\mathbf{W_k}}]_{ii} + [\Sigma_{\mathbf{W_k}}]_{jj}\right)}} + 1\right)$$
```
end for
```

5. EXPERIMENTAL RESULTS

The proposed method has been implemented and tested on a number of test-case grids. Not having access to power grids from industrial designs, and because we need a large number of grids to test our approach under different conditions, we have opted to *generate* a number of grids ourselves. The grid generation process is automatic, and employs a random number generator, as well as user-specified technology and topology parameters. Starting with a square uniform grid of a given size, we proceed to randomly delete a user-specified percentage of nodes, thus rendering the grid structurally non-uniform. Typical geometric and physical grid characteristics (e.g. grid dimensions) as well as characteristics of the fabrication process (e.g. sheet resistance of a particular level of metallization) are given by the user, leading to an initial value of the conductance of every branch. When a node is deleted, the conductances of the remaining surrounding edges (branches) are increased by a random amount around a user-specified percentage of their initial values. The rationale behind this is to allow the non-uniform grid to be loaded with currents comparable to its uniform predecessor while exhibiting comparable IR-drops. The number of V_{dd} (C4) sites and leakage current sources are supplied by the user; the C4s and current sources are then distributed at random over the grid nodes. This implementation was done by writing C programs, and factorization of the grid system was carried out by LU, using routines from
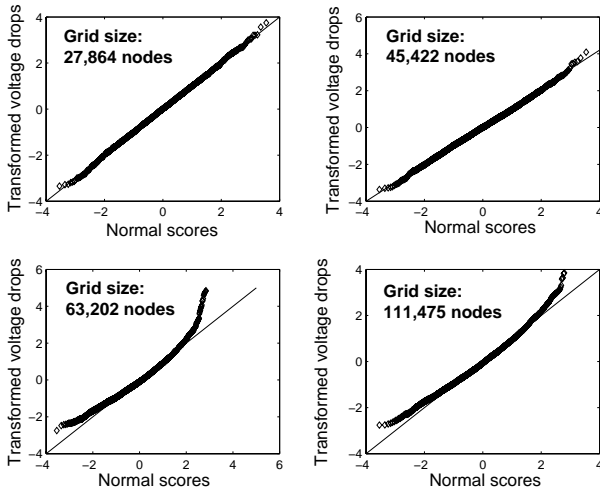
Figure 5: Graphical check of the goodness-of-fit of the voltage drop data against a lognormal distribution.
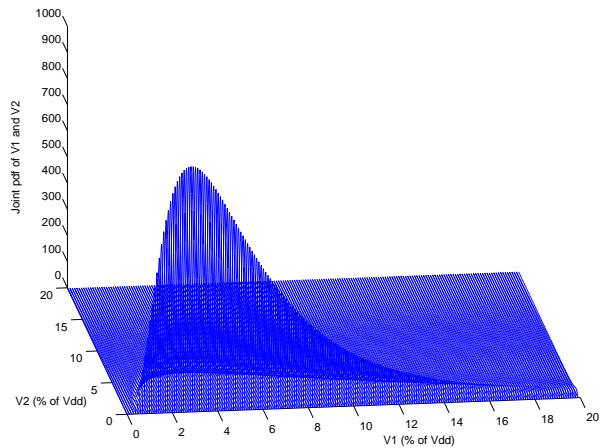

Figure 6: Joint pdf of the voltage drop of two nodes on a power grid of 10,000 nodes

Sparse 1.3 [18].

Fig. 4 plots the entries of the covariance matrix estimated by our column sampling method, versus the exact covariance matrix entries, obtained by full solution of (18). As can be seen from this figure, the estimated (co)variances match well with the exact ones.

Fig. 5 corroborates the fact that voltage drops are well modeled by a lognormal distribution. If indeed voltage drops are lognormally distributed, then their logarithms are normal. We checked this conclusion empirically by generating several grids and collecting voltage drop data, then verifying graphically the goodness-of-fit of their logarithms on normal scores plots [28]. It is clear that voltage drops showed good fits, except for certain outliers, validating the choice of lognormal distributions to model the voltage drops on the power grid, induced by independent, within-die leakage current variations.

Fig. 6 plots the joint bivariate pdf of two randomly selected nodes of a grid of 10,000, after estimating its lognormal parameters. It can be observed from this figure that the probability distribution has very high density in a relatively small area in the voltage drop domain, and vanishes rapidly outside this area.

Table 1 illustrates results where a number of nodes critical nodes were randomly selected to be critical, and the probability of timing violations was calculated using our method. In this table, we compare the proposed column-sampling approach for the estimation of the covariance matrix of critical nodes ("Sampling"), with the result of calculating the entries of this matrix by full solution of (18) ("Solution"). The numerical integration was performed using Monte Carlo techniques [9]. The basic implementation is to estimate the average of the integrand over the domain of integration via random sampling. For numerical integrations in multiple dimensions, Monte Carlo methods are often the only viable alternative. However, when the integrand is strongly concentrated in a relatively small "area" of the domain of integration, with vast differences among values inside and outside that "area" of concentration, sampling the integrand over the entire domain of integration in order to estimate its average may be unwieldy. We were able to obtain converging values for integrals of up to ten dimensions, as shown in table 1, where we often had to break down the domain of integration into multiple smaller subsets. As can be seen, the overall accuracy is satisfactory, while the speedup was up to 130X.

We implemented our technique for grids with a larger number of critical nodes to obtain the mean and covariance matrix of the voltage drop distribution, as shown in table 2. As can be seen, the statistical distributions of grids of up to 78,000 nodes were obtained in less than two minutes. We note that having a large number of critical nodes on the grid is not a typical situation in practice. Short clock periods and a small number of logic levels, characteristic of high-speed circuits, effectively limit the number of such nodes to be considered simultaneously. Furthermore, it will be possible to eliminate neighboring nodes from consideration in a critical path as these nodes display essentially the same voltage levels at all times and constitute practically electrically identical points. In case the number of critical nodes remains high, then more advanced integration methods can be used to calculate the probability of timing violations [12].

6. CONCLUSION

As the impact of process variations and leakage current increases, the power grid may suffer high voltage drops causing failure to meet timing requirements and jeopardizing the design functionality. This work focused on the impact of leakage current variations on the power grid and proposed a statistical model for the grid voltage drops that must be used to account for leakage variations, in order to check the susceptibility of a circuit to timing violations.

7. REFERENCES

[1] Power grid verification. Cadence White Paper, 2002. Online available at: `http://www.cadence.com/whitepapers/4101_PowerGridVerif_WP.pdf`.

[2] R. Ahmadi and F. N. Najm. Timing analysis in presence of power supply and ground voltage variations. In *IEEE/ACM International Conference on Computer-Aided Design*, pages 176–183, San Jose, CA, November 9-13 2003.

[3] N. C. Beaulieu, A. A. Abu-Dayya, and P. J. McLane. Estimating the distribution of a sum of independent lognormal random variables. *IEEE Transactions on Communications*, 43(12):2869–2873, Dec. 1995.

[4] D. Boning and S. Nassif. Models of process variations in device and interconnect. In A. Chandrakasan, W. J.

Table 1: Results on grids with a small number of critical nodes, where calculating the probability of timing violations was possible

Size (#nodes)	# critical nodes	Allowable voltage drop (%V_{dd})	Time (sampling)	Likelihood of timing violations (sampling)	Time (solution)	Likelihood of timing violations (solution)
3,600	5	5	0.9 sec.	28%	9 sec.	33%
6,400	3	10	1.0 sec.	40%	31 sec.	35%
9,900	6	10	1.1 sec.	18%	48 sec.	23%
12,100	8	2	12 sec.	14%	6 min.	12%
25,500	10	1	19 sec.	46%	41 min.	49%

Table 2: Runtimes for mean and covariance matrix estimation on larger grids

Size (#nodes)	# critical nodes	Time (sampling)
33,850	20	30 sec.
44,100	25	43 sec.
57,600	25	61 sec.
67,600	45	75 sec.
78,400	50	98 sec.

Bowhill, and F. Fox, editors, *Design of High-Performance Microprocessor Circuits*, chapter 6. IEEE Press, New York, NY, 2001.

[5] K. A. Bowman, S. G. Duvall, and J. D. Meindl. Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution. In *International Solid State Circuits Conference*, pages 278–279, 2001.

[6] V. De and S. Borkar. Technology and design challenges for low power and high performance. In *ACM/IEEE International Symposium on Low Power Electronics and Design*, pages 163–168, San Diego, CA, August 16-17 1999.

[7] S. G. Duvall. Statistical circuit modeling and optimization. In *International Workshop on Statistical Metrology*, pages 56–63, Honululu, HI, June 2002.

[8] M. Eisele, J. Berthold, D. Schmitt-Landsiedel, and R. Mahnkopf. The impact of intra-die device parameter variations on path delays and on the design for yield of low voltage digital circuits. *IEEE Transactions on Very Large Scale Integrated (VLSI) Systems*, 5(4):360–368, December 1997.

[9] M. Evans and T. Swartz. *Approximating Integrals Via Monte Carlo and Deterministic Methods*. Oxford University Press, Oxford, United Kingdom, 2000.

[10] I. A. Ferzli and F. N. Najm. Statistical estimation of leakage-induced power grid voltage drop considering within-die process variations. In *ACM/IEEE Design Automation Conference*, pages 856–859, Anaheim, CA, June 2-6 2003.

[11] I. A. Ferzli and F. N. Najm. Statistical verification of power grids considering process-induced leakage current variations. In *IEEE/ACM International Conference on Computer-Aided Design*, pages 770–777, San Jose, CA, November 9-13 2003.

[12] J. A. G. Jess, K. Kalafala, S. R. Naidu, R. H. J. M. Otten, and C. Visweswariah. Statistical timing for parametric yield prediction of digital integrated circuits. In *ACM/IEEE Design Automation Conference*, pages 932–937, Anaheim, CA, June 2-6 2003.

[13] T. Kam, S. Rawat, D. Kirkpatrick, R. Roy, G. S. Spirakis, N. Sherwani, and C. Peterson. EDA challenges facing future microprocessor design. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 19(12):1498–1506, December 2000.

[14] T. Karnik, S. Borkar, and V. De. Sub-90nm technologies - challenges and opportunities for CAD. In *IEEE/ACM International Conference on Computer-Aided Design*, pages 203–206, San Jose, CA, November 10-14 2002.

[15] M. Kishor and J. P. de Gyvez. Threshold voltage and power-supply tolerance of CMOS logic design families. In *IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, pages 329–357, Oct. 2000.

[16] C. Kleiber and S. Kotz. *Statistical Size Distributions in Economics and Actuarial Sciences*. John Wiley & Sons, Inc., Hoboken, NJ, 2003.

[17] H. Kriplani, F. N. Najm, and I. Hajj. Pattern independent maximum current estimation in power and ground buses of CMOS VLSI circuits: algorithms, signal correlations, and their resolution. *IEEE Transactions on Computer-Aided Design*, 14(8):998–1012, August 1995.

[18] K. S. Kundert and A. Sangiovanni-Vincentelli. *Sparse User's Guide: A Sparse Linear Equation Solver (Version 1.3a)*. University of California, Berkeley, CA, April 1988. Package and documentation online available at: `http://www.netlib.org/sparse`.

[19] S. Narendra, V. De, S. Borkar, D. Antoniadis, and A. Chandrakasan. Full-chip sub-threshold leakage power prediction model for sub-0.18um CMOS. In *ACM/IEEE International Symposium on Low Power Electronics and Design*, pages 19–23, Monterey, CA, August 12-14 2002.

[20] S. R. Nassif. Design for variability in DSM technologies. In *First International Symposium on Quality Electronic Design*, pages 451–454, 2000.

[21] S. Pant, D. Blaauw, S. Sundareswaran, V. Zolotov, and R. Panda. Vectorless analysis of supply noise induced delay variation. In *IEEE/ACM International Conference on Computer-Aided Design*, pages 184–191, San Jose, CA, November 9-13 2003.

[22] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, NY, 2nd edition, 1984.

[23] L. T. Pillage, R. A. Rohrer, and C. Visweswariah. *Electronic Circuit and System Simulation Methods*. McGraw-Hill, New York, NY, 1995.

[24] R. Rao, A. Srivastava, D. Blaauw, and D. Sylvester. Statistical analysis of subthreshold leakage current for VLSI circuits. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 12(2):131–139, February 2004.

[25] M. S. Srivastava. *Methods of Multivariate Statistics*. Wiley-Interscience, New York, NY, 2002.

[26] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S. Lo, G. A. Sai-Halasz, R. A. Viswanathan, H. C. Wann, S. J. Wind, and H. Wong. CMOS scaling into the nanometer regime. *Proceedings of the IEEE*, 85(4):486–504, April 1997.

[27] S. K. Thompson. *Sampling*. John Wiley & Sons, Inc., New York, NY, 2nd edition, 2002.

[28] R. E. Walpole, R. H. Meyers, and S. L. Meyers. *Probability and Statistics for Engineers and Scientists*. Prentice Hall International, Inc., Upper Saddle River, NJ, 6th edition, 1998.