Redundancy-Aware Power Grid Electromigration Checking Under Workload Uncertainties

Sandeep Chatterjee, *Student Member, IEEE*, Mohammad Fawaz, *Student Member, IEEE*, and Farid N. Najm, *Fellow, IEEE*

Abstract—Electromigration (EM) in on-die metal lines is becoming a significant problem in modern integrated circuits technology. Due to the high levels of current density on the die, the large number of metal lines, and the inherent conservatism in classical full-chip EM models, designers are finding it very hard to meet the area and design specs while guaranteeing EM reliability. The EM problem is most significant in power grid lines, because unlike signal and clock lines, they do not benefit from healing due to their mostly unidirectional currents. In this paper, we develop a new model, referred to as the mesh model, for power grid EM checking which takes into account the inherent redundancy of its mesh structure while determining the reliability. To implement the mesh model, we also develop a framework to estimate the change in statistics of an interconnect as its effective-EM current varies. In order to overcome the conservative assumptions that designers usually make about chip workloads, we also propose a novel vectorless mesh model technique to estimate the average minimum time-to-failure of a power grid under workload uncertainties. The results indicate that the series model, which is currently used in the industry, gives a pessimistic estimate of power grid MTF and reliability by a factor of 3–4. Finally, we exploit multithreading and grid locality to speedup our implementation by almost 6×.

Index Terms—Electromigration (EM), optimization, power grid, reliability, verification.

I. INTRODUCTION

POWER grid verification has become a crucial step in modern integrated circuits (ICs) design, as the robustness of a chip highly depends on the proper functionality of its power grid. A reliable power grid is one that can deliver the required voltage levels to every logic block in the underlying circuit, and that can continue to do so for a certain number of years before failing. Recently, electromigration (EM), a long-term failure mechanism that affects metal lines, has reemerged as a significant concern in very large-scale integration. What is worrying is that the existing EM checking tools are producing overly conservative results due to the series system assumption they make about the grid, as we will explain later.

Moreover, current densities are rising sharply due to the reduction in both supply voltage and width of metal lines [1], causing a loss in the safety margins between the actual and the predicted EM stress. Also, verifying today's power grids has become very expensive due to their large number of nodes. As a result, engineers are forced to reconsider the traditional approaches and to look for new techniques that will allow them to predict the EM stress efficiently with lower conservatism.

Early approaches for EM verification compared worst-case interconnect average current per unit width to a conservative fixed limit to determine whether a line is reliable or not. Statistical electromigration budgeting (SEB) was later introduced in [2] as a way to link the reliability of the metal structure as a whole to the reliability of individual lines. In SEB, the chip is treated as a series system, i.e., a system that is deemed to have failed as soon as any of its components fails. Under such model, and with some simplifying assumptions, the failure rate of a system is the sum of failure rates of its individual components. SEB was first applied to the Alpha 21164 microprocessor and became a standard technique in many industrial Computer Aided Design (CAD) tools. SEB was appealing because it is simple to use and allows some components to have high failure rates as long as the overall chip failure rate is acceptable.

Nevertheless, power grids in modern ICs have a mesh topology rather than the traditional "comb" structure, in the sense that they have many parallel paths between any two given nodes. This implies some sort of redundancy in the design. As a result, power grids are close to a parallel system, and so have a longer lifetime, compared to a series system. In other words, a grid is not necessarily failed if one of its metal lines fails. In many cases, a grid can continue to deliver acceptable voltage levels to every block in the underlying logic circuit even when one or more of its metal lines fail. This observation has always been ignored in both academia and industry, hence the existing EM checking tools do not recognize the benefits of redundancy and always assume that a power grid behaves as a series system. As our results will show, accounting for this redundancy increases the predicted lifetime of a power grid significantly. In some cases, we notice that a grid can tolerate up to 40 or more line failures before it truly fails.

On the other hand, EM degradation in any given metal line depends on the current density through the line. Thus, full-chip reliability depends on the patterns of current drawn by the underlying logic circuitry. Generally, it is unrealistic to expect the user to specify these patterns as this would

require extensive simulation of the chip for millions of clocks cycles at a low enough level of detail so as to produce exact current waveforms. In addition, one might need to verify the grid early in the design flow, when the full details of the underlying logic are not yet available. Ideally, and to overcome the uncertainties about the currents, a vectorless approach is needed, where only a minimal amount of information is known about the current patterns such as bounds on their averages.

In this paper, we develop a new model referred to as the mesh model, that intelligently takes into account the redundancy in the power grid to compute its mean time-to-failure (MTF) and its reliability. We propose a novel approach to estimate the change in statistics of grid lines as their average current density changes in time steps that are comparable to their lifetime. We also develop an exact approach to update the voltage drops in the grid as its metal lines start to fail. The mesh model is then extended to a vectorless framework where the worst-case reliability of the power grid is computed given a set of local and global constraints on the currents drawn from grid. Our data show an increase in the predicted lifetime of 3–4 × compared to the existing series system-based techniques. This implies bigger margins between the predicted and the actual EM stress, making it easier for the designers to sign off on chip designs. Finally, we improve the run-time of the proposed approaches by utilizing: 1) the locality of the power grids and 2) the inherent parallelism present in our algorithms. On an average, our final locality-based multithreaded approach is 6.1× faster than the original approach (which were proposed in preliminary versions of this paper [3], [4]).

The remainder of this paper is organized as follows. In Section II, we present background material on EM, the power grid model, and the Monte Carlo random sampling approach. Section III develops the mesh model and Section IV describes the proposed approach for updating the time-to-failure (TTF) statistics of a metal line in the scenario of changing current densities. Section V describes the vectorless mesh model framework, while Section VI discusses the implementation details. Finally, Section VII shows our experimental results and Section VIII concludes this paper.

II. BACKGROUND

A. Electromigration

EM is a long-term failure mechanism that affects metal lines and vias under high current densities. The force exerted by the flow of electrons can cause the movement of metal atoms in the direction of electron flow. This causes the creation of voids and hillocks. A void is created due to depletion while a hillock is created due to accumulation. Voids generally lead to open circuits or unacceptable resistance increase in a line, whereas protrusion of hillocks usually cause short-circuits between adjacent lines and interlevel conductors at positions where conductors cross over each other or in devices where there are two spaced layers, such as a capacitor. Hillocks due to EM can also result in thin dielectrics, which are further susceptible to dielectric breakdown [5]. In this paper, we assume, for simplicity, that all interconnect failure due to EM are caused by nucleation and/or growth of voids.

Quantitatively, the failure time of a metal line due to EM is usually modeled by a random variable (RV) $\mathbf{T}$ because degradation rates depend on the microstructure of the wire which varies due to random manufacturing variations. Several models were developed to describe the behavior of $\mathbf{T}$ but the simplest and most practical one was empirically developed by Black [6]. According to Black, $\mathbf{T}$ follows a lognormal (LN) distribution, i.e., its logarithm has a normal (Gaussian) distribution. The MTF is given by Black's equation [5], [6]

$$\text{MTF} = \mu = \frac{a}{A} J^{-\eta} \exp\left(\frac{E_a}{kT_k}\right) \tag{1}$$

where A is an experimental constant that depends on the physical properties of the metal line, a is the cross sectional area of the line, J is the effective current density, $\eta > 0$ is the current exponent that depends on the material of the wire and the failure stage, k is the Boltzmann's constant, T_k is the temperature in Kelvin, and E_a is the activation energy for EM. The standard deviation $\sigma_{\ln}$ of $\ln \mathbf{T}$ is usually determined experimentally for a given metal technology. In this paper, we assume that $\sigma_{\ln}$ is the same for all conductors made of a given material. We also assume that the value of T_k is known and is constant across the grid.

For sufficiently short lines, the back stress developed due to accumulation of atoms at the ends of a line can overcome the build up of the critical stress required for creation of a void in the line. In other words, a reversed migration process can occur due to the accumulation of atoms, and this reduces or even compensates the effective material flow toward the anode [5]. For this reason, short lines generally have longer lifetimes and in many cases can be considered immortal; this is called the Blech effect [7]. The Blech effect is quantified is terms of a critical value of the product of current density (J) and length of a line (L), denoted by β_c. This threshold value is useful in circuit design as it determines whether a line is immortal or not as follows: given a line ℓ of length L_ℓ, subject to a current density J_ℓ, then ℓ is considered EM-immune (i.e., immortal) if $J_\ell L_\ell < \beta_c$ and EM-susceptible if $J_\ell L_\ell \geq \beta_c$.

B. Power Grid Model

Because EM is a long-term cumulative failure mechanism, the changes in the current waveforms on short time scales are not very significant for EM degradation. In fact, the standard approach to check for EM failure is to derive a constant effective-EM current from the time varying current waveform [8]. The value thus obtained represents the dc current that effectively gives the same lifetime as the original waveform under identical conditions. Power grid lines mostly carry unidirectional currents, for which effective-EM currents are the same as average branch currents. Also, a power grid is a linear system, which means that the average branch currents can be found by subjecting the grid to average source currents. Therefore, in order to perform EM analysis, it is sufficient to consider a dc model of the grid subject to average source currents.

Let the power grid consist of $n + q$ nodes, where nodes $1 \ldots n$ have no voltage sources attached, and the remaining nodes connect to ideal voltage sources to represent the connections to the external power supply, and let node 0 represent

the ground node. Define I to be the $n \times 1$ vector of all the average source currents such that the entry corresponding to a node with no current source attached is set to zero. Applying modified nodal analysis to the grid leads to

$$G(t)V(t) = I \qquad (2)$$

where $G(t)$ is the conductance matrix of the grid, and $V(t)$ is the vector of voltage drops. As long as the grid is connected, the matrix $G(t)$ is known to be a diagonally dominant symmetric positive definite $\mathcal{M}$-matrix, so that $G^{-1}(t)$ exists and $G^{-1}(t) \geq 0$ [9]. Let b be the number of branches in the grid, $I_{b,l}(t)$ represent the branch currents where $l \in \{1, \ldots, b\}$, and $I_b(t)$ be the vector of all branch currents.

Relating all the branch currents to the voltage drops $V(t)$ across them we get

$$I_b(t) = -R(t)^{-1}M(t)^T V(t) = -R(t)^{-1}M(t)^T G^{-1}(t)I \quad (3)$$

where $R(t)$ is a $b \times b$ diagonal matrix of the branch resistance values and $M(t)$ is an $n \times b$ incidence matrix whose elements are ± 1 or 0 such that the term ± 1 occurs in location m_{kl} of the matrix where node k is connected to the lth branch, else a 0 occurs. The signs of the nonzero terms depend on the node under consideration. If the reference direction for the current is away from the node, then the sign is positive, else it is negative. Please note that the time dependence in $G(t)$, $V(t)$, $I_b(t)$, $R(t)$, and $M(t)$ is solely introduced due to the change in the conductance matrix, that varies over time as grid lines fail due to EM.

C. Mean Estimation by Random Sampling

Consider a continuous RV $\mathbf{x}$ with a certain distribution whose mean is to be estimated by random sampling. We will first discuss the case where $\mathbf{x}$ is normally distributed, and then extend the discussion to the case where $\mathbf{x}$ has an unknown distribution.

Suppose we are sampling from a normal distribution whose variance is unknown. If the true mean of the distribution is μ and its true variance is σ, and if the arithmetic mean of the samples is $\bar{x}_w$ after w iteration, then in order to ensure an upper bound ϵ on the relative error between $\bar{x}_w$ and μ with a confidence of $(1 - \alpha) \times 100\%$, the number of samples w needed is given by [10]

$$w \geq \left(\frac{z_{\alpha/2} s_w}{|\bar{x}_w|\epsilon/(1-\epsilon)} \right)^2 \qquad (4)$$

where s_w is the unbiased estimator of σ and $z_{\alpha/2}$ is the $(1 - \alpha/2)$-percentile of the RV $[(\bar{x}_w - \mu)/(\sigma/\sqrt{w})]$ having a standard normal distribution. The usage of s_w instead of σ (which is unknown) in (4), is acceptable when w is large ($w \geq 30$ as suggested in [10]) because the RV $[(\bar{x}_w - \mu)/(s_w/\sqrt{w})]$ has the t-distribution which approaches the standard normal for large w.

In the general case where the distribution is unknown (not necessarily normal), the RV $[(\bar{x}_w - \mu)/(s_w/\sqrt{w})]$ has been shown to have a distribution that is fairly close to a t-distribution. As before, this t-distribution approaches the standard normal for large w ($w \geq 30$). With this, one can compute the same stopping criterion in (4), which we use throughout this paper as a stopping criterion for Monte Carlo whenever needed.

III. VECTOR-BASED MESH MODEL

As mentioned earlier, traditional methods for EM reliability estimation employ the series system model. A series system is deemed to fail when any of its components fail, i.e., it is only as strong as its weakest link. Given the mesh structure of modern power grids, it is overly conservative to employ the series system for EM checking. The mesh structure of the grid allows multiple paths between any two nodes so that the power grid is not necessarily failed if one of its metal lines fails. In this section, we develop the mesh model and outline the procedure to estimate the MTF and the survival probability of a power grid when the currents drawn from it are known exactly.

Circuit timing is tightly coupled to the node voltage drops [11]. Thus, the integrity of the power grid is evaluated based on how well the supply voltage v_{dd} is conducted to the grid nodes. In other words, for a grid to function as intended, the voltage drop at each node should be smaller than a certain threshold; otherwise soft errors in the underlying logic may occur. A node is said to be safe when its voltage drop meets the corresponding threshold condition, and unsafe otherwise. Let V_{th} be the vector of all the threshold values, which is typically user specified. We assume that $V_{th} > 0$ to avoid trivial cases.

We assume that at $t = 0$, the grid is connected, so that there is a resistive path from any node to another that does not go through a v_{dd} or ground node. Also, we assume that the grid is safe at $t = 0$, i.e., the voltage drops at all the nodes are below their corresponding threshold: $V(0) = G^{-1}(0)I \leq V_{th}$. Notice that if this assumption is not true, then the grid would be unsafe at the production time.

As we move forward in time, the EM-susceptible lines start to fail due to EM. Accordingly, the conductance matrix $G(t)$ of the grid changes and so does $V(t)$. The grid is deemed to fail at the earliest time for which the condition $V(t) = G^{-1}(t)I \leq V_{th}$ is no longer true, which happens when any node in the grid nodes becomes unsafe. This new model is referred to as the mesh model, and will be used to determine the failure time of the grid.

In this paper, we assume that the resistance of a line becomes infinite (an open circuit) when it fails, i.e., we assume that the failure is not gradual but abrupt. This infinite resistance model leads to a simpler and conservative analysis because in reality, a line is said to have failed once its resistance has risen above some threshold. Thus, lines continue to conduct current even after failure, but with higher resistance. Hence, employing the infinite resistance model means we are assuming that a failing line is more degraded than it actually is.

A. MTF and Survival Probability Estimation

Let $\mathbf{T}_m$ be the RV denoting the TTF of the grid according to the mesh model. In order to estimate the MTF ($E[\mathbf{T}_m]$) of the power grid using the mesh model, we perform Monte Carlo analysis. In every iteration, we generate one sample of the grid TTF using the mesh model; we stop once the Monte Carlo convergence criterion (4) is met.

Because I is known, one can find the branch currents in the grid using (3), and then find the JL-product of every line. This allows filtering out the EM-immune lines. The MTF of all the other lines can then be found using Black's equation. For every Monte Carlo iteration, we choose TTF samples for all the EM-susceptible lines from their corresponding LN distributions. We keep failing the resistors in increasing order of their TTF samples until the condition $V(t) \leq V_{th}$ is no longer true. This gives a single grid TTF sample.

We also use Monte Carlo random sampling to estimate the survival probability of a grid up to $\mathcal{Y}$ years, i.e., $\mathcal{P}\{\mathbf{T}_m > \mathcal{Y}\}$. For each Monte Carlo iteration, we obtain a grid TTF sample using the procedure outlined in the previous paragraph, and test whether the grid has survived up to $t = \mathcal{Y}$ or not. Because this represents a Bernoulli trial, we use the bounds derived in [12] to determine how many trials are needed to have an error bound ϵ and a confidence level $(1 - \alpha) \times 100\%$. If w trials were needed, and if the grid was found to be safe at $t = \mathcal{Y}$ in x of those trials, then $\mathcal{P}\{\mathbf{T}_m > \mathcal{Y}\} \approx x/w$.

B. Generating Time-to-Failure Samples

As mentioned before, branch currents are needed to discover the EM-immune lines, and to find the MTF of all the other lines using Black's equation. Since the grid will be changing over time due to the failure of its components, the branch currents will also change. For now, we will focus on generating TTF samples at $t = 0$. The issue of updating the TTF samples with changing currents will be detailed in Section IV.

If G_0 is the conductance matrix of the original grid [i.e., $G_0 = G(0)$], then the vector of initial voltage drops can be written as $V_0 = V(0) = G_0^{-1}I$. This allows writing

$$I_b(0) = I_b = -R(0)^{-1}M(0)^T G_0^{-1}I. \tag{5}$$

Consider a line l with cross sectional area a_l. At $t = 0$, let $I_{b,l}$ be the branch current flowing in it. Then, if line l is not EM-immune, its MTF μ_l should be computed using Black's equation (1) and can be rewritten as

$$\mu_l = \frac{a_l^{\eta+1}}{A} |I_{b,l}|^{-\eta} \exp\left(\frac{E_a}{kT_k}\right) \tag{6}$$

where we used $J_l = |I_{b,l}|/a_l$. We know that TTF of line l is modeled by a LN RV $\mathbf{T}_l$. However, for Monte Carlo analysis, a TTF sample τ_l should be assigned to every EM-susceptible line at the start of every iteration. This can be done by sampling a real number ψ_l from the standard normal distribution $\mathcal{N}(0, 1)$, and then applying the transformation below [10]

$$\tau_l = \mu_l \exp\left(\psi_l \sigma_{\ln} - 0.5\sigma_{\ln}^2\right) \tag{7}$$

where σ_{ln} is the standard deviation of $\ln \mathbf{T}_l$ and is assumed to be constant for a given material [3]. If γ_l^T is the row of $-R(0)^{-1}M(0)^T G_0^{-1}$ that corresponds to line l, then $I_{b,l} = \gamma_l^T I$ due to (5), and hence, given a sample ψ_l from the standard normal distribution, we can find a sample TTF τ_l for every line l, using (6) and (7) as follows:

$$\tau_l = \frac{a_l^{\eta+1}}{A} |\gamma_l^T I|^{-\eta} \exp\left(\frac{E_a}{kT_k}\right) \exp\left(\psi_l \sigma_{\ln} - 0.5\sigma_{\ln}^2\right). \tag{8}$$

C. Computing Voltage Drops

Checking if the grid has failed at a particular point in time requires checking the condition $V(t) \leq V_{th}$. Because the infinite resistance model is used, $V(t)$ changes only when a line fails, and remains the same between any two consecutive line failures. Therefore, $V(t)$ should be recomputed every time a line fails. One way of doing that is by updating $G(t)$ followed by its LU factorization and forward/backward solves to obtain $V(t) = G^{-1}(t)I$. Unfortunately, performing an LU factorization, from scratch, every time a line fails is expensive. However, since we are modeling the failure of every line by an open circuit, we can write the change in G corresponding to the kth line failure as a rank-1 matrix $-\Delta G_k$. This corresponds to the removal of a conductance (connected between nodes x and y with $x > y$) from the conductance matrix by reversing the element stamping procedure for that particular conductance. Accordingly, ΔG_k can be written as $\Delta G_k = u_k u_k^T$ with

$$u_k = \sqrt{g_k}(e_x - e_y)$$

where e_λ is a column vector of appropriate size containing 1 at the λth location and zeros at all other locations with e_0 being a vector of all zeros and g_k is the conductance of the line that is being removed.

After the failure of k lines, let U_k be the $n \times k$ matrix such that

$$U_k = \begin{bmatrix} u_1 & u_2 & \dots & u_k \end{bmatrix}.$$

Therefore, $U_k U_k^T = \sum_{j=1}^{k} u_j u_j^T = \sum_{j=1}^{k} \Delta G_j$. This means we can write the vector of voltage drops V_k after the failure of k lines as

$$V_k = \left(G_0 - \sum_{j=1}^{k} \Delta G_j\right)^{-1} I = \left(G_0 - U_k U_k^T\right)^{-1}I. \tag{9}$$

1) Sherman–Morrison–Woodbury Formula: Given the equation above and the initial vector of voltage drops V_0, it is possible to obtain V_k efficiently (i.e., without computing the inverse of $G_0 - U_k U_k^T$) by using the Sherman–Morrison–Woodbury formula [13]. In essence, the formula asserts that the inverse of a rank-k correction of some invertible matrix can be computed by doing a rank-k correction to the inverse of the original matrix. The formula is also known as the matrix inversion lemma, and allows writing the following:

$$\left(G_0 - U_k U_k^T\right)^{-1} = G_0^{-1} + G_0^{-1}U_k(I_k - U_k^T G_0^{-1}U_k)^{-1}U_k^T G_0^{-1} \tag{10}$$

where I_k is the $k \times k$ identity matrix. This assumes that G_0 is nonsingular (which we know because the grid is assumed to be connected and safe at $t = 0$), and that $W_k \triangleq I_k - U_k^T G_0^{-1}U_k$ is also nonsingular. We will first handle the case where W_k is nonsingular, and discuss the singularity case later on. Using (9) and (10), we have

$$V_k = G_0^{-1}I + \left[G_0^{-1}U_k\left(I_k - U_k^T G_0^{-1}U_k\right)^{-1}U^T G_0^{-1}\right]I. \tag{11}$$

Define $Z_k = G_0^{-1}U_k = [G_0^{-1}u_1 \ldots G_0^{-1}u_k]$. Because $G_0^{-1}I = V_0$, we can finally write

$$V_k = V_0 + Z_k W_k^{-1} y_k \qquad (12)$$

where $y_k = U_k^T V_0$. The vector V_k must be computed using (12) for every $k \in \{1, 2, \ldots\}$ until the condition $V_k \le V_{th}$ is no longer true. Computing V_0 should be done only once by doing an *LU* factorization of G_0 and a forward/backward solve. For every k, Z_k must be updated by appending the column vector $G_0^{-1}u_k$, which can be easily computed using forward/backward substitutions. Finally, the inverse of the dense $k \times k$ matrix W_k must be computed. If k is small, we can factorize W_k for every k in $\mathcal{O}(k^3)$ time. However, k can be quite large for big grids, and hence computing the *LU* factorization of W_k becomes expensive as k increases. To overcome this limitation, we propose a further refinement based on the Banachiewicz–Schur form so that the complexity of finding W_k^{-1} is reduced to $\mathcal{O}(k^2)$.

2) Banachiewicz–Schur Form: Let $d_k = 1 - u_k^T G_0^{-1} u_k$ and $b_k = [-u_1^T G_0^{-1} u_k \ldots -u_{k-1}^T G_0^{-1} u_k]^T$. One can prove, using the Banachiewicz–Schur form [14], that W_k^{-1} can be expressed it terms of W_{k-1}^{-1} as follows:

$$W_k^{-1} = \begin{bmatrix} W_{k-1}^{-1} + \dfrac{W_{k-1}^{-1} b_k b_k^T W_{k-1}^{-1}}{s_k} & -\dfrac{W_{k-1}^{-1} b_k}{s_k} \\ -\dfrac{b_k^T W_{k-1}^{-1}}{s_k} & \dfrac{1}{s_k} \end{bmatrix} \qquad (13)$$

where $s_k \triangleq d_k - b_k^T W_{k-1}^{-1} b_k$. Moreover, one can prove that, if $a_k \triangleq (b_k^T x_{k-1} + u_k^T v_0)/s_k$, then

$$x_k = \begin{bmatrix} x_{k-1} + a_k W_{k-1}^{-1} b_k \\ -a_k \end{bmatrix} \qquad (14)$$

where $x_k \triangleq W_k^{-1} y_k$ and $x_{k-1} \triangleq W_{k-1}^{-1} y_{k-1}$. Details of deriving (13) and (14) can be found in [3]. We can use (13) and (14) to directly update W_k^{-1} and x_k from their previous values. Notice that W_k^{-1} is required because, in the next iteration, $W_k^{-1} b_{k+1}$ is needed to compute x_{k+1} using (14). The implementation requires a single matrix–vector product ($\mathcal{O}(k^2)$) and $\mathcal{O}(k^2)$ additions and divisions.

3) Case of Singularity: One can prove that $G_0 - U_k U_k^T$ is invertible if and only if W_k is invertible [13], and that W_k is invertible if and only if $s_k \neq 0$ [14]. Therefore, if for some k, s_k is found to be zero, then we know that W_k is singular and hence $G_0 - U_k U_k^T$ is also singular. In this particular case, V_k cannot be computed. Physically, a grid has a singular conductance matrix when a subset of its nodes becomes disconnected from all the voltage sources. This could happen because we are modeling the failure of a line by an open circuit. Overall, the grid is deemed to fail at the earliest time for which the condition $V(t) \le V_{th}$ is no longer true or when $s_k = 0$.

IV. ACCOUNTING FOR CHANGES IN BRANCH CURRENTS

An interconnect failure in the power grid changes the currents through all the surviving interconnects and hence affects their residual lifetime. Because we have adopted an abrupt model of line failure, the currents will experience step changes
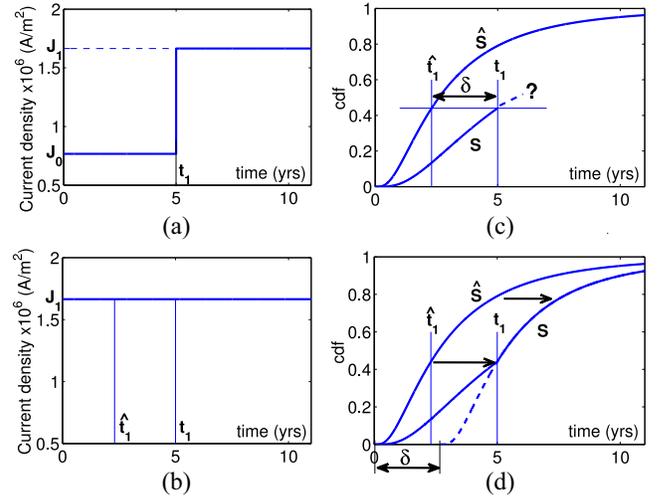
Fig. 1. Proposed approach for single-step case. Current density profile for (a) S and (b) $\hat{S}$. (c) What is the CDF of S for $t > t_1$. (d) Proposed solution.

over time, which we assume remain uni-directional. We now describe a novel approach to estimate the change in failure statistics of an interconnect when its effective current density changes over time.

A. Motivation—The Single Step Case

Consider a thought experiment in which a large set S of N isolated conductors are tested for their failure times. The testing starts at $t = 0$. Let the current densities through all the conductors be identical and given by the following step function:

$$J(t) = \begin{cases} J_0, & 0 \le t \le t_1 \\ J_1, & t_1 < t < \infty \end{cases} \qquad (15)$$

where $J_0 \neq J_1$ and t_1 is large, such that many conductors may have failed before t_1. This current profile is shown in Fig. 1(a). The population S is fresh at $t = 0$, but as time progresses, it suffers damage due to EM and conductors start failing. We are interested in determining the distribution of the RV that describes the statistics of the TTFs of the population.

Unfortunately, the effective-EM current model [8] is not applicable to this case because it implicitly assumes that the resulting effective-EM current density is applied to all the conductors throughout their lifetime. It is meant to handle current waveforms that change on a much smaller time scale as compared to the lifetime of the conductors; whereas in our experiment the current changes on a time scale that is comparable to the TTFs of the conductors. Hence, many conductors might have failed exclusively due to J_0. This motivates the need for a new approach to estimate the statistics of the surviving subpopulation.

To motivate such an approach, consider another set $\hat{S}$ of N conductors identical to those in S. Suppose $\hat{S}$ is subjected to a current density of J_1 for all $t \ge 0$, as shown in Fig. 1(b). Let $F(t)$ be the cumulative distribution function (CDF) of the population S and $\hat{F}(t)$ be the CDF of $\hat{S}$. Clearly, $\hat{F}(t)$ is known to be LN. Define $\hat{t}_1$ to be such that $\hat{F}(\hat{t}_1) = F(t_1)$ as shown in Fig. 1(c). The difference $\delta = t_1 - \hat{t}_1$ is easy to compute, as we will demonstrate later. For now, we focus on the key

question: "what is the failure distribution of S after t_1?" As already pointed out, traditional EM work is not helpful here. We now provide a proposal for answering this question.

Considering the two populations: 1) S at time t_1 and 2) $\hat{S}$ at time $\hat{t}_1$, notice the following.

1) The two populations started out fresh with the same number of conductors, and the expected number of surviving members of the two populations are exactly the same, due to the fact that $\hat{F}(\hat{t}_1) = F(t_1)$. Therefore, the two populations have experienced an identical level of deterioration.

2) The two populations are subjected to exactly the same current stress J_1, as they move forward in time, i.e., $t_1 + x$ and $\hat{t}_1 + x$, with $x \geq 0$ for S and $\hat{S}$, respectively.

Therefore, we expect that, going forward in time, both populations will see the same instantaneous failure rate, i.e., $\lambda(t_1 + x) = \hat{\lambda}(\hat{t}_1 + x)$, $\forall x \geq 0$, or

$$\lambda(t) = \hat{\lambda}(t - \delta) \quad \forall t \geq t_1. \tag{16}$$

Since $\hat{\lambda}(t)$ is the failure rate of a LN distribution, it follows that the failure rate of the surviving subpopulation of S, i.e., $\lambda(t_1 + x)$, is that of a LN. Thus, we propose that the statistics of the surviving population of S for $t > t_1$ be obtained by shifting the origin of the LN that gives rise to $\hat{\lambda}(t)$ by δ so that the continuity of $F(t)$ at $t = t_1$ is maintained, as shown in Fig. 1(d). The mean of the shifted LN distribution is equal to the mean of the LN that gave rise to $\hat{\lambda}(t)$. This proposal can be generalized as follows. At any time $t > 0$, the TTF statistics for the surviving population of S is described by a (section of) a shifted LN distribution, the mean of which (relative to its start time) is given by Black's equation, with J being the current density at time t. As we will see later, this generalization allows us to calculate TTF statistics for multiple current changes. Also, if there are no current changes, the proposed approach gracefully falls back to the use of a single LN.

1) Determining $F(t)$ and δ: We define two RVs $\mathbf{T}_0$ and $\mathbf{T}_1$, where $\mathbf{T}_0$ describes the TTF distribution of S when it is subjected to J_0 $\forall t \geq 0$, and $\mathbf{T}_1$ describes the TTF distribution when S is subjected to zero current density for $t \leq \delta$ and J_1 for $t > \delta$. Clearly, the CDFs $F_{T_0}(t)$ and $F_{T_1}(t - \delta)$ of $\mathbf{T}_0$ and $\mathbf{T}_1$ are known to be LN. Note that the LN distribution of $\mathbf{T}_1$ is shifted and originates at $t = \delta$. Following the arguments of the previous paragraph, we propose that the CDF of S can be written as:

$$F(t) = \begin{cases} F_{T_0}(t), & 0 \leq t \leq t_1 \\ F_{T_1}(t - \delta), & t_1 < t < \infty \end{cases} \tag{17}$$

where the time shift $\delta \in (-\infty, t_1)$ is found using the continuity constraint $F_{T_1}(t_1 - \delta) = F_{T_0}(t_1)$ or

$$\Phi\left[\frac{\ln(t_1 - \delta) - \mu_{\ln,1}}{\sigma_{\ln,1}\sqrt{2}}\right] = \Phi\left[\frac{\ln t_1 - \mu_{\ln,0}}{\sigma_{\ln,0}\sqrt{2}}\right] \tag{18}$$

where Φ the is standard normal CDF, $\mu_{\ln,k} = E[\ln \mathbf{T}_k]$, and $\sigma_{\ln,k}^2 = \text{Var}(\ln \mathbf{T}_k)$. Also, define $\mu_{T,k} = E[\mathbf{T}_k]$. Using the

properties of LN distribution, we can write the following relation:

$$\mu_{\ln,k} = \ln(\mu_{T,k}) - 0.5\sigma_{\ln,k}^2. \tag{19}$$

Since Φ is monotonic, we equate the terms in the brackets of (18) and use (19) to obtain

$$\ln\left(\frac{t_1 - \delta}{\mu_{T,1}}\right) = \frac{\sigma_{\ln,1}}{\sigma_{\ln,0}} \ln\left(\frac{t_1}{\mu_{T,0}}\right) + 0.5\sigma_{\ln,1}(\sigma_{\ln,0} - \sigma_{\ln,1}). \tag{20}$$

If the dependence of $\sigma_{\ln}$ with regard to the damage accumulated due to EM is empirically known beforehand, we could have used it to solve (20) for δ. However, since we are not aware of any such relationship, we assume that the value of $\sigma_{\ln}$ for S at time t is the same as that of the fresh population at $t = 0$ (i.e., $\sigma_{\ln,0} = \sigma_{\ln,1}$). Also, from Black's equation, we know that $(\mu_{T,1}/\mu_{T,0}) = (J_0/J_1)^\eta$, and hence the relative time shift δ between $\mathbf{T}_1$ and $\mathbf{T}_0$ is

$$\delta = t_1\left[1 - \left(\frac{J_0}{J_1}\right)^\eta\right]. \tag{21}$$

We next obtain the CDF in the general case of multiple change in currents.

B. Case of Multiple Change in Currents

Consider a second thought experiment with S in which the current density profile is given as

$$J(t) = J_k, \quad t_k < t \leq t_{k+1}, \quad k = 0, 1 \ldots n \tag{22}$$

where $J_{k-1} \neq J_k$ $\forall k > 0$, $t_0 = 0$, and $t_{n+1} = \infty$. It is interesting to note that (22) is the typical current density profile of a surviving interconnect in the power grid, where the kth failing interconnect has TTF $\tau = t_k$.

As per the previous discussion, for each time span $t_k < t \leq t_{k+1}$, the statistics is described by a RV $\mathbf{T}_k$ that has a LN distribution originating at some $t = \Delta_k$. The mean $\mu_{T,k}$, for each time span $t_k < t \leq t_{k+1}$, is given by Black's equation with $J = J_k$. In order to satisfy the continuity constraint, each RV $\mathbf{T}_k$ has a time shift of δ_k with respect to $\mathbf{T}_{k-1}$, with $\mathbf{T}_0$ having a shift of $\delta_0 = 0$. This implies that the distribution for $\mathbf{T}_k$ originates at $\Delta_k = \sum_{i=0}^{k} \delta_i$, with $\Delta_0 = \delta_0 = 0$. The CDF of S can now be written as

$$F(t) = F_{T_k}(t - \Delta_k), \quad t_k < t \leq t_{k+1}, \quad k = 0, 1 \ldots n. \tag{23}$$

The time shift δ_k between $\mathbf{T}_k$ and $\mathbf{T}_{k-1}$ can be found using the continuity constraint

$$\Phi\left[\frac{\ln(t_k - \Delta_k) - \mu_{\ln,k}}{\sigma_{\ln,k}\sqrt{2}}\right] = \Phi\left[\frac{\ln(t_k - \Delta_{k-1}) - \mu_{\ln,k-1}}{\sigma_{\ln,k-1}\sqrt{2}}\right].$$

Using $\Delta_k = \Delta_{k-1} + \delta_k$, (19) and the assumption that $\sigma_{\ln}$ does not change when the current changes (i.e., $\sigma_{\ln,0} = \ldots = \sigma_{\ln,n} = \sigma_{\ln}$), we have

$$\ln\left(\frac{t_k - \Delta_{k-1} - \delta_k}{\mu_{T,k}}\right) = \ln\left(\frac{t_k - \Delta_{k-1}}{\mu_{T,k-1}}\right).$$
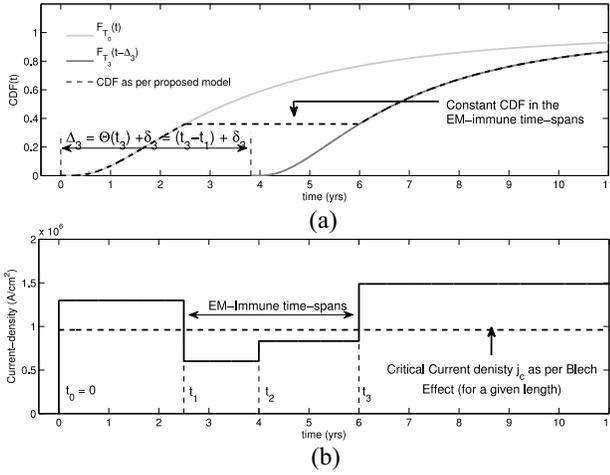
Fig. 2. Proposed approach for multi-step case, with EM-immune time spans in between. (a) CDF plot for a conductor undergoing current changes, with EM-immune time spans in between. (b) Current density profile through the conductor.

By equating the terms in brackets, this can be simplified to

$$\delta_k = (t_k - \Delta_{k-1})(1 - r_k) = \left(t_k - \sum_{i=1}^{k-1} \delta_i\right)(1 - r_k) \quad (24)$$

where $r_k = (\mu_{T,k}/\mu_{T,k-1}) = (J_{k-1}/J_k)^\eta$.

C. Incorporating Blech Effect

The previous analysis assumed $J_k L > \beta_c \; \forall k$. However, due to change in current density, we might have EM-immune and EM-susceptible time spans interspersed with each other. Let M and B be the set of integers k, where k denotes the time span $t_k < t \le t_{k+1}$, so that $M = \{k : J_k L \le \beta_c\}$ and $B = \{k : J_k L > \beta_c\}$. Clearly, $M \cap B = \emptyset$ and $M \cup B$ is the entire time period.

The extension of the above framework to incorporate the Blech effect is primarily based on the observation that a surviving conductor cannot fail for $k \in M$. Accordingly, the corresponding probability of failure is zero and the associated CDF is a constant function, as shown in Fig. 2. The RV $\mathbf{T}_k$, that has a shifted LN distribution originating at $t = \Delta_k$, exists only for $k \in B$.

The definitions of δ_k and Δ_k are altered slightly to account for Blech effect. Consider a general scenario in which $(p-1)$ consecutive EM-immune time spans are sandwiched between two EM-susceptible time spans. To be precise, $k - p, k \in B$ and $k - p + 1, \ldots k - 1 \in M$. Then, δ_k is time shift of $\mathbf{T}_k$ relative to $\mathbf{T}_{k-p}$ needed to maintain the continuity constraint if the in-between EM-immune time spans are removed

$$\delta_k = (t_{k-p+1} - \Delta_{k-p})\left(1 - \left(\frac{J_{k-p}}{J_k}\right)^\eta\right). \quad (25)$$

Note that (25) reduces to (24) for $p = 1$. Also

$$\Delta_k = \Theta(t_k) + \sum_{i=0, i \in B}^{k} \delta_i \quad (26)$$

where $\Theta(t_k)$ is sum of all EM-immune time spans up to t_k.

D. Updating TTF Sample

Consider a conductor $\mathcal{C}$ of the set S subjected to the current density profile (22). Clearly, the TTF of $\mathcal{C}$ changes because the RV describing the statistics of the population changes for each time span. Assume, for the sake of argument, that $\mathcal{C}$ survives for $t > t_k$ and $0, k \in B$. At $t_0(= 0)$, $\mathcal{C}$ has a TTF given by [using (7)]

$$\tau_0 = \mu_{T,0}. \exp\left(\psi \sigma_{\ln} - 0.5\sigma_{\ln}^2\right) \quad (27)$$

where the symbols are as defined before. At $t = t_k$, when the kth current change occurs, the TTF of $\mathcal{C}$ is updated using the following relation:

$$\tau_k = \Delta_k + \mu_{T,k} \exp\left(\psi \sigma_{\ln} - 0.5\sigma_{\ln}^2\right) \quad (28)$$

where ψ is the same sample value from Φ as used in (27). The offset Δ_k is added so that τ_k is referred from $t = 0$. For $k \in M$, τ_k is defined to be ∞.

Theorem 1: Consider a conductor having the current density profile of (22). Let $k - p, k \in B$, and $k - p + 1, \ldots k - 1 \in M$. Then, if (25) and (28) are used to find the offset (δ_k) and TTF (τ_k) for the conductor, we always have

$$\tau_k = t_k + (\tau_{k-p} - t_{k-p+1})\left(\frac{J_{k-p}}{J_k}\right)^\eta \quad (29)$$

so that $\tau_k > t_k$.

The proof of this result can be found in [3].

E. Selective Updates

In estimating the MTF and survival probability using the mesh model, a bulk of the computation effort (nearly 50%) is spent on updating the TTFs of the surviving lines. Completely removing TTF updates from the mesh model in order to speedup computation is not a feasible solution as: 1) it results in significant errors (up to 40%) and 2) the average speedup obtained is only about 1.2× due to slower Monte Carlo convergence [15]. In this section, we present an approach that is aimed at reducing the CPU time spent in updating the TTFs in a way that minimally affects the Monte Carlo convergence. Another advantage of this approach is that it gives us a knob to trade-off accuracy for speed, if required.

For any given line, let $V_{b,k}$ represent its line voltage drop after the kth current change. We can rewrite (29) in terms of line voltage drops as (for simplicity, we use $p = 1$)

$$\begin{aligned} \tau_k &= t_k + (\tau_{k-1} - t_k)\left(V_{b,k-1}/V_{b,k}\right)^\eta \\ &= t_k + (\tau_{k-1} - t_k)\left((V_{b,k} - \Delta V_{b,k})/V_{b,k}\right)^\eta \end{aligned} \quad (30)$$

where $\Delta V_{b,k}$ is the change in line voltage drop, i.e., $V_{b,k} = V_{b,k-1} + \Delta V_{b,k}$. Clearly, $\Delta V_{b,k} \to 0 \implies \tau_k \to \tau_{k-1}$. Hence, we can choose not to update the TTF for lines where change in the line voltage drop is less than some user-specified threshold $\Delta V_b^\star$. This qualitatively gives us a knob to trade-off accuracy for speed. Also, it was observed that a smaller value of $\Delta V_b^\star$ gives better Monte Carlo convergence, with $\Delta V_b^\star = 0$ being the same as the original approach. Let $\pi(\mathsf{R})$ represent the set of grid lines that should be updated (i.e., those lines where the change in line voltage drop is greater than $\Delta V_b^\star$)

after failure of some interconnect R. Determining $\pi(R)$ exactly is still a CPU intensive task. Instead, we now propose a simple heuristic that conservatively estimates $\pi(R)$.

When an interconnect in the power grid fails, not all the nodes are equally impacted. In fact, only a subset of the nodes (and the associated lines), in the immediate vicinity of the failing line R are significantly impacted. This locality can be exploited to estimate $\pi(R)$. Let $\mathcal{N} = \{1, 2, \ldots, n\}$ denote the set of all the nodes in the power grid. Also, let $\partial V_{[k]}/\partial R$ denote the change in the voltage drop of node k with respect to the failure of interconnect R. Define the set $\mathcal{N}_R$ as follows:

$$\mathcal{N}_R = \left\{ k \in \mathcal{N} : \frac{\partial V_{[k]}}{\partial R} > \frac{\Delta V_b^\star}{2} \right\}. \tag{31}$$

Finding $\mathcal{N}_R$ can be swiftly done by checking which nodes in the grid presented a change in their voltage drop larger than $\Delta V_b^\star/2$ after the failure of R. As a conservative approximation, we consider any line connected to a node $k \in \mathcal{N}_R$ to be in the set $\pi(R)$. As we will see later, using selective updates gives huge speedups with minimal loss in accuracy.

V. Vectorless Mesh Model

The analysis in the previous sections assumed that the currents drawn from the power grid are known exactly. Unfortunately, as we mentioned earlier, it is not always realistic to expect users to specify precise values of the current sources or the power dissipation of each block in the underlying circuit. These values change depending on the activity of the blocks, thus producing a large variety of possible current waveforms that can be drawn from the grid. In addition, grid design and verification cannot wait until the chip design is complete, and is typically done early in the design flow where the details of the different blocks are not fully known. Therefore, a vectorless approach is needed to capture the uncertainty about the current waveforms and to assess the reliability of the grid over all the different operation scenarios in the underlying circuit. In this section, we show how to extend the vector-based mesh model to a constraints-based framework that only requires limited information about the underlying logic.

A. Modal Probabilities

Modern ICs have complex multimodal behavior, where major blocks of the chip have different modes of operation (such as stand-by, low power, high performance, etc.). Specifying the block power dissipation requires knowledge of how often these modes are exercised. For every circuit block j, let $k \in \{1, \ldots, r\}$ enumerate the different modes of operation and I_{jk} denote the block average supply current in that mode. The overall average supply current of that block is given by $I_j = \sum_{k=1}^{r} \alpha_{jk} I_{jk}$, where $0 \leq \alpha_{jk} \leq 1$ represent the probability of being in different modes with the constraint that $\sum_{k=1}^{r} \alpha_{jk} = 1$. We propose that it is reasonable to expect the user to specify the currents I_{jk} using the average power dissipation of each block in every power mode. The mode probabilities α_{jk} are generally harder to assess, but users are expected to be able to specify values for some of them, or

narrow ranges for others. If α denotes the $nr \times 1$ vector of all the mode probabilities (considering all the n blocks connected to the n grid nodes, having r modes of operation each), then we can write

$$\alpha_{\min} \leq \alpha \leq \alpha_{\max} \tag{32}$$

where $\alpha_{\min}$ and $\alpha_{\max}$ have entries between 0 and 1, and contain any information the user may have about the modes of operation.

The user can also specify bounds on the average current of every block, if available. This allows us to infer other constraints on α in the form

$$I_{\ell,\min} \leq L\alpha \leq I_{\ell,\max} \tag{33}$$

where L is an $n \times nr$ matrix such that $I = L\alpha$. The matrix L contains information about the currents drawn by the circuit blocks in each power mode.

Since chip components rarely draw their maximum currents simultaneously, global constraints are also used. For instance, if a certain limit is specified on the average power dissipation of the chip, then one may say that the sum of all the current sources is no more than a certain upper bound. In general, the same concept can be applied for groups of current sources forming functional blocks with known upper and lower bounds on their average power [16]. If m is the total number of global constraints, then we can write

$$I_{g,\min} \leq SL\alpha \leq I_{g,\max} \tag{34}$$

where S is an $m \times n$ matrix that only contains 0s and 1s and indicates which current sources are present in each global constraint. The matrix contains a 1 at the kth entry of the ith row if the kth circuit block (current sources) is present in the ith global constraint.

One last set of constraints should be added to guarantee that $\sum_{k=1}^{r} \alpha_{jk} = 1$ for every block j

$$B\alpha = \mathbf{1}_n \tag{35}$$

where B is an $n \times nr$ matrix containing only 1s and 0s such that the vector $B\alpha$ contains the sum of mode probabilities per block in each of its entries, and $\mathbf{1}_n$ is a vector of size n containing only 1s. Together, all the constraints presented above define a feasible space of mode probabilities, denoted by $\mathcal{F}_\alpha$, such that $\alpha \in \mathcal{F}_\alpha$ if and only if, α satisfies (32)–(35).

For example, consider a circuit having three blocks with two modes of operation each: high performance and low power. Assume that the blocks draw, respectively, 0.2, 0.3, and 0.25 A on average in high performance mode, and 0.1, 0.2, and 0.15 A in low power mode. Also, let α_{11}, α_{21}, and α_{31} denote the probabilities of the blocks being in high performance mode, and α_{12}, α_{22}, and α_{32} the probabilities of being in low power mode. If the average currents of the blocks are I_1, I_2, and I_3, and if $I = \begin{bmatrix} I_1 & I_2 & I_3 \end{bmatrix}^T$ and $\alpha = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{21} & \alpha_{22} & \alpha_{31} & \alpha_{32} \end{bmatrix}^T$ then we can write

$$I = L\alpha = \begin{bmatrix} 0.2 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0.15 \end{bmatrix} \alpha.$$

The following is a possible set of constraints that a user can specify:

$$\begin{bmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.3 \\ 0.6 \\ 0.1 \end{bmatrix} \leq \alpha \leq \begin{bmatrix} 0.7 \\ 0.6 \\ 0.5 \\ 0.9 \\ 0.9 \\ 0.9 \end{bmatrix} \quad \begin{bmatrix} 0.11 \\ 0.21 \\ 0.17 \end{bmatrix} \leq L\alpha \leq \begin{bmatrix} 0.18 \\ 0.29 \\ 0.24 \end{bmatrix}$$

$$\begin{bmatrix} 0.35 \\ 0.4 \end{bmatrix} \leq SL\alpha = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} L\alpha \leq \begin{bmatrix} 0.41 \\ 0.48 \end{bmatrix}$$

$$B\alpha = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \alpha = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

For every feasible setting of α, the overall block average currents are different, and the reliability of the power grid is correspondingly different. Our goal is to look for the worst-case reliability of the power grid given all the possible feasible combinations of α.

B. Current Feasible Space

As a first step toward evaluating the worst-case reliability of the grid, we transform the feasible space $\mathcal{F}_\alpha$ to the current domain. This helps reduce the number of variables from nr to n, as well as the number of constraints. It is easy to see that replacing $L\alpha$ by I in (33) and (34) results in the first set of constraints defining the feasible space of currents

$$I_{\ell,\min} \leq I \leq I_{\ell,\max} \qquad (36)$$
$$I_{g,\min} \leq SI \leq I_{g,\max}. \qquad (37)$$

On the other hand, given the constraints on the individual α's for every current source, we can find lower and upper bounds for all the sources, as follows. Recall that every current source I_j can be written as $I_j = \sum_{k=1}^{r} \alpha_{jk} I_{jk}$, and let α_j denote the vector of size r of all the mode probabilities corresponding to I_j, then due to (32) we can write

$$\alpha_{j,\min} \leq \alpha_j \leq \alpha_{j,\max}$$

where $\alpha_{j,\min}$ and $\alpha_{j,\max}$ contain the upper and lower bounds on the entries of α_j as specified in (32). Due to (35), we can write: $\sum_{k=1}^{r} \alpha_{jk} = 1$, and hence, we can find bounds $I_{j,\min}$ and $I_{j,\max}$ on I_j by solving the following two linear programs (LP):

$$\text{Min/Max} \quad \sum_{k=1}^{r} \alpha_{jk} I_{jk}$$
$$\text{subject to} \quad \alpha_{j,\min} \leq \alpha_j \leq \alpha_{j,\max}$$
$$\sum_{k=1}^{r} \alpha_{jk} = 1. \qquad (38)$$

The LPs above should be solved for every current source in the power grid. If any of the LPs turns out to be infeasible, then the user specifications are not consistent. Notice that due to the structure of the LPs above, we do not need to use any of the classical LP solving methods (simplex or interior point). In fact, the two claims below show how to compute the solutions directly. Assume, without loss of generality, that the modes of

operation of block j are sorted in decreasing order of their power consumption, i.e., $I_{j1} \geq I_{j2} \geq \ldots \geq I_{jr}$. Also, call $\alpha_{jk,\min}$ and $\alpha_{jk,\max}$, $k \in \{1, \ldots, r\}$, the entries of the vectors $\alpha_{j,\min}$ and $\alpha_{j,\max}$, respectively.

Claim 1: Consider the largest $h \leq r$ for which $\sum_{k=1}^{h-1} \alpha_{jk,\max} \leq 1$. Then, the solution to the maximization problem in (38) is

$$\alpha_{jk} = \begin{cases} \alpha_{jk,\max} & \text{for } k = 1, \ldots, h-1 \\ 1 - \sum_{k=1}^{h-1} \alpha_{jk} & \text{for } k = h \\ \alpha_{jk,\min} & \text{for } k = h+1, \ldots, r. \end{cases}$$

Proof: To see why this works, notice that the problem is infeasible if $\sum_{k=1}^{r} \alpha_{jk,\min} > 1$ or $\sum_{k=1}^{r} \alpha_{jk,\max} < 1$. Assuming that the problem is feasible, we notice that we can replace the last equality constraint by the inequality constraint $\sum_{k=1}^{r} \alpha_{jk} \leq 1$ without changing the optimal solution. The reason is that if we were able to fit all the α's without reaching equality, then $\sum_{k=1}^{r} \alpha_{jk,\max} < 1$, making the original problem infeasible, which contradicts our assumption. Accordingly, we want to show that the greedy approach explained above solves the problem below

$$\text{Maximize} \quad \sum_{k=1}^{r} \alpha_{jk} I_{jk}$$
$$\text{subject to} \quad \alpha_{j,\min} \leq \alpha_j \leq \alpha_{j,\max}$$
$$\sum_{k=1}^{r} \alpha_{jk} \leq 1.$$

Consider the following change of variables $\forall k \in \{1, \ldots, r\}$:

$$w_k = \frac{\alpha_{jk} - \alpha_{jk,\min}}{\alpha_{jk,\max} - \alpha_{jk,\min}}.$$

In the space of w, and $\forall k \in \{1, \ldots, r\}$, the problem becomes

$$\text{Maximize} \quad \sum_{k=1}^{r} c_k w_k + \sum_{k=1}^{r} I_{jk} \alpha_{jk,\min}$$
$$\text{subject to} \quad 0 \leq w_k \leq 1$$
$$\sum_{k=1}^{r} b_k w_k \leq d \qquad (39)$$

where $c_k = I_{jk}(\alpha_{jk,\max} - \alpha_{jk,\min})$, $b_k = (\alpha_{jk,\max} - \alpha_{jk,\min})$, and $d = 1 - \sum_{k=1}^{r} \alpha_{jk,\min}$. Because the original problem is assumed to be feasible, we have $d \geq 0$. Also, we notice that $c_k \geq 0$ and $b_k \geq 0$ for every k. Ignoring the constant term $\sum_{k=1}^{r} I_{jk} \alpha_{jk,\min}$ in the objective function, (39) becomes an LP relaxation of the well known 0-1 Knapsack problem [17] for which the optimal solution can be found using a greedy approach. If $c_1/b_1 \geq c_2/b_2 \geq \ldots \geq c_r/b_r$ (which is true because $c_k/b_k = I_{jk}$ and the I_{jk}'s are assumed to be sorted in this order), then the optimal solution can be found as follows: set $w_1 = w_2 = \ldots = w_{h-1} = 1$, $w_h = d - \sum_{k=1}^{h-1} b_k$, and $w_{h+2} = \ldots = w_r = 0$, where $h \leq r$ is the largest possible such that $\sum_{k=1}^{h-1} w_k \leq d$. Transforming this solution back into the α space gives the solution described earlier. ∎

Claim 2: Consider the smallest $g \geq 1$ for which $\sum_{k=g+1}^{r} \alpha_{jk,\max} \leq 1$. Then, the solution to the minimization

problem in (38) is

$$\alpha_{jk} = \begin{cases} \alpha_{jk,\max} & \text{for } k = g+1, \ldots, r \\ 1 - \sum_{k=g+1}^{r} \alpha_{jk} & \text{for } k = g \\ \alpha_{jk,\min} & \text{for } k = 1, \ldots, g-1. \end{cases}$$

The proof the claim 2 is almost identical to the proof of claim 1, and is omitted.

Ultimately if all the LPs turn out to be feasible, we obtain a lower and an upper bound on every current source. However, (36) also provides similar bounds, hence, all the bounds should be combined to obtain

$$I_{\min} \leq I \leq I_{\max}. \tag{40}$$

Overall, we obtain a new feasible space of currents, that we call $\mathcal{F}$, such that $I \in \mathcal{F}$ if and only if, I satisfies (40) and (37).

Back to the example in the previous section, the resulting reduced set of constraints in the current domain would be

$$\begin{bmatrix} 0.14 \\ 0.22 \\ 0.21 \end{bmatrix} \leq I \leq \begin{bmatrix} 0.17 \\ 0.25 \\ 0.24 \end{bmatrix}$$

$$\begin{bmatrix} 0.35 \\ 0.4 \end{bmatrix} \leq SI = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} I \leq \begin{bmatrix} 0.41 \\ 0.48 \end{bmatrix}.$$

Now, our goal becomes to look for the worst-case reliability of the grid given all the possible feasible combinations of I. For that, we find the average minimum TTF of the grid subject to $I \in \mathcal{F}$. We do that by performing a Monte Carlo analysis as before. In every iteration, we choose a sample from the standard normal distribution for every line in the grid, and we find the smallest grid TTF that can be obtained using the mesh model given any $I \in \mathcal{F}$, and the set of samples chosen for the lines. Recall that these samples are used to sample failure times for the lines using (8) which, in this case, yields an expression for every TTF since I is not fixed. An exact solution to this problem was proposed in [4], but it was computationally very expensive. Here, we propose an approximate solution that is based on simulated annealing (SA) and is very fast in practice ($\sim 100\times$ faster for a 586 node grid [4]). It uses the TTF estimator developed in Section IV.

C. Overview of Simulated Annealing

SA is a random search global optimization techniques that occasionally allows uphill movements. SA has been used to solve both discrete and continuous optimization problems. In this paper, we are concerned with continuous optimization since our feasible space is a convex polytope of currents, i.e., we are concerned with problems of the form

$$f^* = \min_{x \in X} f(x) \tag{41}$$

where $X \subseteq \mathbb{R}^n$ is a continuous compact domain. SA randomly generates a candidate point at every iteration and decides whether to move to it through a random mechanism based on a parameter called temperature. In order to define a complete SA algorithm, one should appropriately define how to select the next candidate point, how and when to accept the next candidate point, how to update the temperature, and when to converge.

Accepting the next candidate point is typically done randomly based on an acceptance function $\mathcal{A}$. Here, we use the Metropolis function, the most widely used acceptance function

$$\mathcal{A}(x_k, y_{k+1}, T_k) = \min\left\{1, \exp\left(-\frac{f(y_{k+1}) - f(x_k)}{T_k}\right)\right\}$$

where x_k is the current point, y_{k+1} is the candidate point, and T_k is the current temperature. The acceptance function always returns a number between 0 and 1 denoting the probability of accepting the point y_{k+1}. In practice, this is done by sampling a number q between 0 and 1 from the uniform distribution and then choosing the next current point x_{k+1} as follows:

$$x_{k+1} = \begin{cases} y_{k+1} & \text{if } q \leq \mathcal{A}(x_k, y_{k+1}, T_k) \\ x_k & \text{otherwise.} \end{cases} \tag{42}$$

Notice that, if $f(y_{k+1}) \leq f(x_k)$, then the acceptance function returns 1, and hence the new current point is $x_{k+1} = y_{k+1}$ (accepted with a probability equal to 1). Otherwise, y_{k+1} is accepted with a probability that depends on T_k and how large is the gap $|f(x_k) - f(y_{k+1})|$. A large gap or a low temperature results in a low acceptance probability. Accepting an ascent step from $f(x_k)$ to $f(y_{k+1})$ is sometimes necessary to avoid being trapped at a local minimum, and is called hill climbing.

Updating the temperature is done based on a cooling schedule $\mathcal{U}$. Here, we use

$$T_k = \mathcal{U}(T_0, k) = a^{\lfloor \frac{k}{M} \rfloor} T_0 \tag{43}$$

where k is the SA iteration index, T_0 is the starting temperature, a is a constant between 0.8 and 0.99, and M is an integer. Notice that this allows the temperature to decrease by the factor a after each group of M iterations. Convergence occurs when the temperature becomes less than some small positive number T_ϵ.

What remains is discussing how the next candidate point inside the feasible space is found. Notice that choosing a starting point x_0 can be done by solving a feasibility problem in X, which in our case can be done using a LP because our feasible space is a convex polytope. Finding the next candidate point y_{k+1} can be done by uniformly generating a random direction θ in space such that $\|\theta\|_2 = 1$, and then computing the set Λ defined as follows:

$$\Lambda = \Lambda(x_k, \theta) = \{\lambda \in \mathbb{R} : x_k + \lambda\theta \in \mathcal{F}\}.$$

Finding Λ can be done easily when X is convex. In fact, it can be shown that, if $x_k \in X$, then the set Λ can be written as

$$\Lambda = \{\lambda \in \mathbb{R} : \lambda_{\min} \leq \lambda \leq \lambda_{\max}\}$$

where $\lambda_{\min} \leq 0$ and $\lambda_{\max} \geq 0$. To choose a new random point, it is enough to uniformly sample a value for λ from the set Λ, and compute the new point accordingly: $y_{k+1} = x_k + \lambda\theta$. More details about how to choose the next candidate point can be found in [18].

VI. IMPLEMENTATION

The overall flow for estimating the MTF of a power grid using the vector-based mesh model is given in the flow chart of Fig. 3. Note that the TTF of the first failing resistor for any
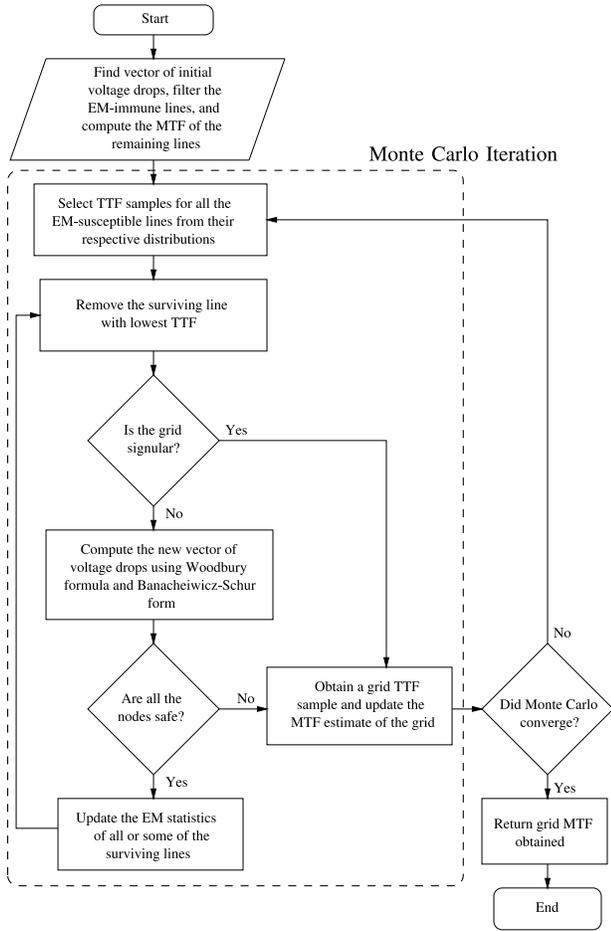
Fig. 3. MTF estimation using vector-based mesh model.

Fig. 4. Average minimum grid TTF estimation using vectorless mesh model.

Monte Carlo iteration gives us the grid TTF as per the series model. Estimating the average minimum TTF of a power grid using the vectorless mesh model is as shown in Fig. 4. Given the structure of the sequential algorithms, it can be easily seen that the Monte Carlo iterations within them can be parallelized. At the beginning of the execution, we use the main thread for initializing the shared power grid data: computing V_0 and finding the MTF of all the EM-susceptible lines at $t = 0$. At the end of the initialization phase, the main thread creates p child threads and goes into a waiting state. Each child thread starts executing one Monte Carlo iteration. Every time a child thread completes an iteration and generates a new grid TTF sample, it updates the shared average and standard deviation and the total number of iterations completed so far. At this point, it checks condition (4) to see if the stopping criteria has been satisfied. If not, it picks up a new set of TTF samples for the metal lines and starts another Monte Carlo iteration. On the other hand, if the stopping criteria is met, the child thread returns. After all the child threads have returned, the control passes back to the main thread which then cleans up the memory and prints the results. The locking scheme and memory management aspects of the parallelization are omitted due to lack of space.

VII. EXPERIMENTAL RESULTS

The algorithms presented in Figs. 3 and 4 were implemented in C++. We used `boost::threads` for parallelization.

Two types of test grids were used for running the experiments: 1) internal and 2) external. Internal grids were synthesized as per user specifications, including grid dimensions, metal layers, pitch, and width per layer. The supply voltages and current sources were randomly placed on the grid. The technology specifications were consistent with $1.1V$ 65 nm CMOS technology. The grids named DC1-DC9 are internal grids. The external grids are part of IBM power grid benchmarks [19]. These grids are dual grids, but we used only the v_{dd} part of the grids. The voltage drop threshold V_{th} was defined to be 10% of v_{dd} for all nodes in a grid. The interconnect material was assumed to be aluminum. As for the EM model employed, we use the Blacks model with an activation energy of 0.9 eV, a current exponent $\eta = 1$ (we assume that the lifetime is dominated by the time taken for the void to grow), a nominal temperature $T_k = 373$ K (the user can provide any temperature profile for the grid lines; we use $T_k = 373$ K as an average temperature throughout the chip), and a critical Blech product $\beta_c = 3000$ A/cm. All the experiments were carried out on a 3.4 GHz Linux machine with 32 GB of RAM. The Monte Carlo constants we used for our simulations were $\alpha = 0.05$ for which $z_{\alpha/2} = 1.96$, $\epsilon = 0.05$ for the algorithm in Fig. 3, and $\epsilon = 0.1$ for the algorithm in Fig. 4.

To assess the reduction in pessimism as a result of using the mesh model, we compare: 1) the power grid MTF (μ_s and μ_m) and 2) the survival probability ($\mathcal{P}_s$ and $\mathcal{P}_m$) as estimated using the series and the mesh model, respectively. Table I compares μ_m and μ_s using a gain ratio (μ_m/μ_s). The gain ratio is dependent on $\Gamma = \max(V_0)$; as Γ decreases, the gain ratio increases. If Γ is close to V_{th}, the mesh model degenerates to the series model. Given a reasonable difference between V_{th} and V_0, the gain ratio is 3–4 for all the grids, with the average gain ratio for the reported grids being 3.83. We also observe that for each Monte Carlo iteration, the largest grid (DC9) requires an average of 125 lines to fail before the grid fails, which shows the inherent redundancy in the grid.

Table I also shows the speedup and accuracy of the selective updates approach, with $\Delta V_b^\star = 10^{-3}V$. The maximum error as compared to the original approach (in which the TTF of

TABLE I
COMPARISON OF POWER GRID MTF ESTIMATED AND CPU TIME USING THE SERIES MODEL,
THE MESH MODEL, AND THE MESH MODEL WITH SELECTIVE UPDATES

Power Grid				Γ^b	Series	Mesh (Original)			Gain Ratio	Mesh with Selective Updates					Error^d
Name	Nodes^a	C4's	Sources		MTF (yrs)	MTF (yrs)	Avg fails	t_{st} (hrs)^c		MTF (yrs)	t_{su} (hrs)^c	Speed-up (t_{st}/t_{su})	t_{mt} (hrs)^c	Speed-up (t_{st}/t_{mt})	(%)
ibmpg5	249K	100	237K	4.80	3.09	9.96	17.7	0.28	**3.23**	9.97	0.06	4.79x	0.04	**7.34x**	0.14
ibmpgnew1	316K	494	179K	9.36	2.88	12.94	58.1	1.53	**4.50**	13.01	0.32	4.81x	0.19	**8.04x**	0.52
ibmpg6	404K	132	381K	5.73	2.74	10.47	30.6	0.73	**3.82**	10.5	0.15	4.96x	0.10	**7.14x**	0.28
ibmpg3	410K	494	100K	6.77	2.39	9.86	47.2	0.90	**4.12**	9.87	0.30	3.05x	0.18	**5.13x**	0.14
DC7	449K	17K	28K	4.57	3.36	12	65.6	1.16	**3.57**	12.13	0.36	3.20x	0.19	**5.94x**	1.07
ibmpg4	475K	312	133K	7.35	2.76	10.28	41.1	1.30	**3.72**	10.30	0.52	2.52x	0.30	**4.39x**	0.17
ibmpgnew2	718K	494	179K	7.72	2.54	10.12	44.9	2.34	**3.98**	10.14	0.77	3.04x	0.44	**5.32x**	0.20
DC8	1M	39K	63K	3.99	2.88	10.79	91.5	3.52	**3.75**	10.81	1.16	3.02x	0.59	**5.93x**	0.15
DC9	1.8M	70K	112K	3.68	2.61	9.97	124.5	8.01	**3.81**	9.98	2.70	2.97x	1.46	**5.48x**	0.05

^{a}Number of nodes in v_{dd} rails after merging the short paths as equivalent node
$^b\Gamma = \max(V_0)$, expressed as $\%v_{dd}$. Note that $V_{th} = 10\%v_{dd}$.
$^c t_{st}$, t_{su} and t_{mt} are the wall times for single threaded original approach, single threaded selective updates approach and multi-threaded selective updates approach, respectively
^{d}Error is reported between the MTFs calculated using the original and the selective updates approach

TABLE II
SURVIVAL PROBABILITY ESTIMATION

Power Grid		Γ ($\%v_{dd}$)	$\mathcal{Y}$ (yrs)	$\mathcal{P}_s$	$\mathcal{P}_m$	Seq. runtime (hrs)	Par. runtime (hrs)	Speed-up
Name	Nodes							
ibmpg5	249K	4.80	4.94	0.035	0.980	0.08	0.05	1.59x
ibmpgnew1	316K	9.36	4.60	0.025	1.000	0.16	0.09	1.67x
ibmpg6	404K	5.73	4.38	0.029	0.996	0.16	0.10	1.62x
ibmpg3	410K	6.77	3.83	0.016	0.998	0.25	0.15	1.72x
DC7	449K	4.57	5.38	0.018	0.998	0.21	0.12	1.77x
ibmpg4	475K	7.35	4.42	0.020	0.998	0.42	0.24	1.76x
ibmpgnew2	718K	7.72	4.07	0.018	0.994	0.49	0.28	1.74x
DC8	1M	3.99	4.61	0.006	0.998	0.53	0.29	1.81x
DC9	1.8M	3.68	4.18	0.006	0.998	1.25	0.66	1.91x

all surviving interconnects are updated after each failure) is 1.07%, and we obtain significant speedups ranging between 2.5–5×. This alone shows the effectiveness of the single-threaded selective updates approach. The run-time is further improved by using four threads to parallelize Monte Carlo iterations. As seen in Table I, by using a multithreaded selective updates approach, the time required to predict MTF for the 1.8 M node grid is brought down from 8 to 1.46 h (5.48× speedup), while the error in estimation is only 0.05%. Overall, for all the reported grids, the multithreaded selective updates approach achieves an average speedup of 6.07× over the single-threaded original approach, with the average error being only 0.30%.

Fig. 5(a) plots run-time versus number of nodes for the vector-based MTF estimation engine. Although we obtain significant speedups, the scalability of the selective updates approach degrades when compared to the original approach. For the multithreaded selective updates approach, the total time taken to estimate the MTF scales as $\sim\mathcal{O}(n^{1.35})$, whereas the original approach scales as $\sim\mathcal{O}(n^{1.32})$. The increase in complexity is primarily attributed to weak metal lines which fail in almost all Monte Carlo iterations. This leads to a lot of data sharing between threads running parallel Monte Carlo iterations and hence results in higher complexity than expected.

Table II compares the survival probability estimated as per the series and the mesh model based on user-specified values for $\mathcal{Y}$. By choosing $\epsilon = \alpha = 0.05$, and by using the bounds derived in [12], the total number of iterations required is 489. The multithreaded variant used four threads for parallelization. From Table II, we can clearly see that by
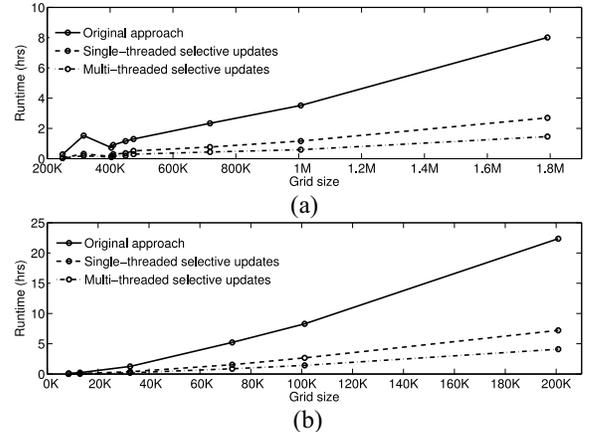
Fig. 5. Run-time analysis of the proposed approaches. (a) Mesh model with known source currents. (b) SA-based vectorless mesh model.

taking redundancies into account, the mesh model consistently predicts a higher survival probability as compared to the series model. Also, it is quite fast, with the multithreaded version requiring only 40 min to estimate the survival probability for the largest grid. The multithreaded approach in each case has only upto 2× speedup, mainly due to the high memory access to computation ratio for sparse matrix operations.

Reliability of a power grid is essentially its survival probability at different points in time. Hence, if the TTF statistics of the power grid following the mesh and series model is described by the RVs $\mathbf{T}_m$ and $\mathbf{T}_s$, then we can empirically determine their probability distribution function (PDF) and

TABLE III
COMPARISON OF POWER GRID AVERAGE MINIMUM TTF AND CPU TIME USING THE VECTORLESS MESH MODEL
AND THE VECTORLESS MESH MODEL WITH SELECTIVE UPDATES

Power Grid				Simulated Annealing		Simulated Annealing with Selective Updates					Error^f
Name	Nodes	C4's	Sources	Average Min TTF (yrs)	t_{vst} (hrs)^e	Average Min TTF (yrs)	t_{vsu} (hrs)^e	Speed-up (t_{vst}/t_{vsu})	t_{vmt} (hrs)^e	Net speed-up (t_{vst}/t_{vmt})	(%)
DC1	8.4K	1.6K	552	10.09	0.10	10.19	0.04	2.44x	0.01	**6.98x**	0.99
DC2	13K	240	812	10.53	0.22	10.61	0.08	2.65x	0.04	**6.27x**	0.76
DC3	32.5K	552	2K	10.09	1.26	10.10	0.38	3.31x	0.21	**6.14x**	0.01
DC4	72.7K	1.3K	4.6K	10.39	5.22	10.47	1.54	3.39x	0.87	**6.00x**	0.77
DC5	101K	1.7K	6.3K	10.28	8.29	10.40	2.67	3.10x	1.43	**5.80x**	1.17
DC6	201K	3.4K	12.7K	10.84	22.36	10.85	7.22	3.10x	4.10	**5.45x**	0.01

^{e} t_{vst}, t_{vsu} and t_{vmt} are the wall times for single threaded original SA approach, single threaded selective updates SA approach and multi-threaded selective updates SA approach, respectively.
^{f} Error is reported between the average minimum MTFs as calculated between the original and the selective updates approach.
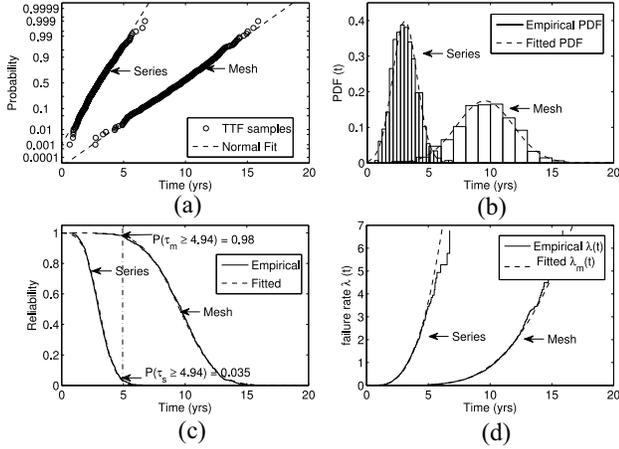
Fig. 6. Complete statistics for IBMPG5 grid (vector-based). (a) Goodness-of-fit. (b) PDF. (c) Reliability. (d) Failure rate.

Fig. 7. Statistics for DC3 grid (vectorless). (a) Goodness-of-fit. (b) PDF. (c) Reliability. (d) Failure rate.

their reliability by calculating the survival probabilities for different values of $\mathcal{Y}$. Following the above idea, a complete statistical analysis was done for the IBMPG5 grid to estimate its RVs $\mathbf{T}_m$ and $\mathbf{T}_s$. A total of 489 grid TTF samples were obtained for each model. Using the goodness-of-fit methods, it was found that the samples could be fitted well by a normal distribution [Fig. 6(a)], which means that the distributions of $\mathbf{T}_m$ and $\mathbf{T}_s$ are close to normal. This is further verified by the plots in Fig. 6(b)–(d) which show a good agreement between the empirical PDF, reliability and the failure rate functions and the actual curves plotted with the parameters obtained from the goodness-of-fit plot. Fig. 6 again shows that the series model is highly pessimistic. Also, note that since $\mathbf{T}_m$ and $\mathbf{T}_s$ are now shown to have a normal distribution, the use of (4) as a stopping criterion to estimate MTF is also verified.

On the other hand, Table III shows the average minimum grid TTF estimated and the required run-time using the vectorless mesh model algorithm of Fig. 4. It also shows that the vectorless approach benefits by incorporating the selective updates heuristic. This is evident from the fact that the maximum error as compared to the original approach is found to be 1.17%, whereas the speedup obtained ranges between 2.4 and 3.4×. Moreover, Table III also shows the speedup obtained by using the multithreaded vectorless approach (using four threads) with selective updates. For the largest grid having 201 K nodes, the total estimation time was reduced from 22.36 to 4.10 h (5.45× speedup), while the error in the estimation is

only 0.01%. Overall, we notice an average speedup of 6.11× over the single-threaded original vectorless approach, and an average error of 0.61%.

Fig. 5(b) shows the plot of run-time versus number of nodes for the vectorless EM checking engine. As before, we observe that, although we obtain significant speedups, the scalability of the original version is better than the multithreaded version. The total time taken by the multithreaded vectorless approach scales as $\sim \mathcal{O}(n^{1.79})$, whereas the original approach scales as $\sim \mathcal{O}(n^{1.72})$.

We did a complete statistical analysis for the DC3 grid to estimate the distribution of its worst-case TTF. A total of 489 minimum grid TTF samples were obtained using the vectorless approach and it was found that the samples could be fitted well by a normal distribution [Fig. 7(a)]. This is further validated by the extremely good agreement shown between the empirical and actual theoretical curves for PDF, reliability and failure rate functions [Fig. 7(b)–(d)]. Therefore, the use of (4) as a stopping criterion to estimate the average minimum grid TTF is also verified for the vectorless engine.

Moreover, our results show that there is a large separation between the average minimum and the average maximum TTF of the grid. The separation was found to vary between 10 and 25 years. This implies that the MTF of the grid is highly sensitive to the change in currents and that it is not enough to compute the TTF of the grid at an arbitrary feasible point.

Lastly, to check how sensitive SA is to the random seed used to travel the feasible space, we minimized the TTF of grid DC6 (using SA with selective updates) 50× while keeping the same TTF samples for the grid lines and while changing the random seed that controls how the next candidate points are being chosen. We obtained 50 different minimums having a mean of 10.65 years and a standard deviation of 0.68 years, leading to a mean to standard deviation ratio of 0.064. This shows that SA is not very sensitive to any change in the random seed, and hence, the result of SA is, up to certain extent, a good estimate of the minimum grid TTF.

VIII. CONCLUSION

With the latest technology scaling and high demand for low voltage designs, EM is becoming increasingly problematic in power grids. We introduced two fast and efficient redundancy-aware early verification techniques for EM checking. The first approach is vector-based and assumes full knowledge of the loading currents while the second approach relaxes this assumptions and follows a constraint-based framework at the cost of a higher run-time. We showed that the mesh model we proposed reduces the pessimism that arises in the state of the art EM checking tools. To make the estimation more efficient, we explored locality within the grid to selectively update the TTF samples of the surviving interconnects whenever an interconnect fails, and we also implemented a multithreaded version of each algorithm. One might argue that the grids above are small (especially the ones used to find the average worst-case TTF) when compared to full-chip grids containing millions of nodes. However, our methods can be applied to the top-level main feeder network of the grid that is not very large and that should be tested early in the design flow, or to small parts of the grid using the proper macromodeling techniques. Collectively, we believe that the techniques presented can fill real and diverse design needs.

REFERENCES

[1] J. Warnock, "Circuit design challenges at the 14 nm technology node," in *Proc. ACM/IEEE 48th Design Autom. Conf. (DAC)*, New York, NY, USA, Jul. 2011, pp. 464–467.

[2] J. Kitchin, "Statistical electromigration budgeting for reliable design and verification in a 300-MHz microprocessor," in *VLSI Circuits Tech. Dig.*, Kyoto, Japan, 1995, pp. 115–116.

[3] S. Chatterjee, M. Fawaz, and F. N. Najm, "Redundancy-aware electromigration checking for mesh power grids," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design*, San Jose, CA, USA, Nov. 2013, pp. 540–547.

[4] M. Fawaz, S. Chatterjee, and F. N. Najm, "A vectorless framework for power grid electromigration checking," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design*, San Jose, CA, USA, Nov. 2013, pp. 553–560.

[5] A. Christou and M. Peckerar, *Electromigration and Electronic Device Degradation*. New York, NY, USA: Wiley, 1994.

[6] J. R. Black, "Electromigration—A brief survey and some recent results," *IEEE Trans. Electron. Devices*, vol. 16, no. 4, pp. 338–347, Apr. 1969.

[7] I. A. Blech, "Electromigration in thin aluminium on titanium nitride," *J. Appl. Phys.*, vol. 47, no. 4, pp. 1203–1208, 1976.

[8] K. Lee, "Electromigration recovery and short lead effect under bipolar- and unipolar-pulse current," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Anaheim, CA, USA, Apr. 2012, pp. 6B.3.1–6B.3.4.

[9] F. N. Najm, *Circuit Simulation*. Hoboken, NJ, USA: Wiley, 2010.

[10] J. E. Freund, I. R. Miller, and R. Johnson, *Probability and Statistics for Engineers*, 6th ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2010.

[11] R. Ahmadi and F. N. Najm, "Timing analysis in presence of power supply and ground voltage variations," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design*, San Jose, CA, USA, Nov. 2003, pp. 176–183.

[12] F. N. Najm, "Statistical estimation of the signal probability in VLSI circuits," Coordinat. Sci. Lab., Univ. Illinois Urbana-Champaign, Champaign, IL, USA, Tech. Rep. UILU-ENG-93-2211, Apr. 1993.

[13] N. J. Higham, *Functions of Matrices: Theory and Computation*, 1st ed. Philadelphia, PA, USA: SIAM, 2008.

[14] Y. Tian and Y. Takane, "Schur complements and Banachiewicz–Schur forms," *Electron. J. Linear Algebra*, vol. 13, pp. 405–418, Dec. 2005.

[15] S. Chatterjee, "Redundancy aware electromigration checking for mesh power grids," Master's thesis, Dept. Electr. Comput. Eng., Univ. Toronto, Toronto, ON, Canada, Aug. 2013.

[16] D. Kouroussis and F. N. Najm, "A static pattern-independent technique for power grid voltage integrity verification," in *Proc. ACM/IEEE Design Autom. Conf.*, Anaheim, CA, USA, Jun. 2003, pp. 99–104.

[17] H. Kellere, U. Pferschy, and D. Pisinger, *Knapsack Problems*. Berlin, Germany: Springer, 2004.

[18] M. Fawaz, "Electromigration reliability analysis of power delivery networks in integrated circuits," Master's thesis, Dept. Electr. Comput. Eng., Univ. Toronto, Toronto, ON, Canada, Aug. 2013.

[19] S. R. Nassif, "Power grid analysis benchmarks," in *Proc. Asia South Pac. Design Autom. Conf.*, Seoul, Korea, 2008, pp. 376–381.

Sandeep Chatterjee (S'14) received the B.Tech. degree in electrical engineering (with high distinction) from the Indian Institute of Technology Roorkee, Roorkee, India, in 2011, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2013, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering.

His current research interests include computer-aided design for integrated circuits with a focus on the reliability, verification, and analysis of power grids.

Mohammad Fawaz (S'08) received the B.E. degree in computer and communications engineering (with high distinction) from the American University of Beirut, Beirut, Lebanon, in 2011, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2013, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering.

His current research interests include computer-aided design for integrated circuits with a focus on the verification, reliability, and analysis of power grids.

Farid N. Najm (S'85–M'89–SM'96–F'03) received the B.E. degree in electrical engineering from the American University of Beirut, Beirut, Lebanon, and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign (UIUC), Urbana, IL, USA, in 1983 and 1989, respectively.

He is a Professor with the Edward S. Rogers, Sr., Department of Electrical and Computer Engineering (ECE), University of Toronto, Toronto, ON, Canada. From 1989 to 1992, he was with Texas Instruments, Dallas, TX, USA. He then joined the Department of ECE, UIUC, as an Assistant Professor and became an Associate Professor in 1997. In 1999, he joined the Department of ECE, University of Toronto, where he is currently a Professor and the Chair. His current research interests include CAD for very large-scale integration, with an emphasis on circuit level issues related to power, timing, variability, and reliability. He has authored the book entitled *Circuit Simulation* (New York, NY, USA: Wiley) in 2010.

Dr. Najm was a recipient of the IEEE Transactions on CAD Best Paper Award, the NSF Research Initiation Award, the NSF CAREER Award, and the Design Automation Conference (DAC) Prolific Author Award. He was an Associate Editor of the IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS from 2001 to 2009 and the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS from 1997 to 2002. He has served on the Executive Committee of the International Symposium on Low-Power Electronics and Design (ISLPED) 1999–2013 and the Technical Committee of various conferences, including IEEE/ACM International Conference for Computer Aided Design, DAC, CICC, International Symposium on Quality Electronic Design, and ISLPED. He has served as the TPC Chair and the General Chair for ISLPED. He is a fellow of the Canadian Academy of Engineering.