A Vectorless Framework for Power Grid Electromigration Checking*

Mohammad Fawaz
Department of ECE
University of Toronto
Toronto, Ontario, Canada
mohammad.fawaz@mail.utoronto.ca

Sandeep Chatterjee
Department of ECE
University of Toronto
Toronto, Ontario, Canada
chatte45@eecg.utoronto.ca

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

ABSTRACT

Electromigration (EM) in the on-die metal lines has re-emerged as a significant concern in modern VLSI circuits. The higher levels of temperature on die and the very large number of metal lines, coupled with the conservatism inherent in traditional EM checking strategies, have led to a situation where trying to guarantee EM reliability often leads to unacceptably conservative designs that may not meet the area or performance specs. Due to unidirectional currents, this problem is most significant in the power and ground grids. Thus, this work is aimed at reducing the pessimism in EM prediction for power/ground grids. There are two sources for the high pessimism: 1) the use of the traditional series model for EM checking and 2) pessimistic assumptions about the chip workload and the corresponding supply currents. To address this problem, we propose a framework for EM checking that allows users to specify conditions-of-use type constraints that help capture realistic chip workload and which includes the use of a novel mesh model for EM prediction in the grid, instead of the traditional series model.

Keywords

Power grid, Electromigration, Verification, Redundancy, Optimization

1. INTRODUCTION

Power grid verification has become an essential step in modern integrated circuits (IC) design. To guarantee the robustness of a chip, test engineers must verify that the power grid can provide the required voltage levels to the underlying logic, and that it can continue to do so for a certain number of years before failing. Electromigration, a long term failure mechanism that affects metal lines, is becoming a significant problem in VLSI especially in the power grid. What is worrying is that the existing tools for power grid EM checking are producing pessimistic results (because they rely on the *series model* as we will explained soon), and hence, the safety margins between the predicted EM stress and the EM design rules are becoming smaller.

Historically, electromigration checking tools relied on *worst-case* current density limits for individual grid lines. Later on, *Statistical Electromigration Budgeting* (SEB) was introduced in [1] in which the *series model* is employed with other simplifying assumptions leading to a simple expression of the failure rate as the sum of failure rates of individual components. This model was applied to the Alpha 21164 microprocessor and became a standard technique in many industrial CAD tools. SEB was appealing because it related the reliability of circuit components to the reliability of the whole system. In addition, SEB is simple to use and allows some components to have high failure rates as long as the sum of all the failure rates is acceptable.

Nonetheless, modern power grids are meshes rather than the traditional "comb" structure. The mesh structure allows multiple paths between any two nodes, and accordingly, the power grid is not necessarily failed if one of its metal lines fails, but it can tolerate more failures until the voltage drops at its nodes become unacceptable. This implies some level of redundancy in the grid, which has traditionally been ignored when the series model was assumed. As a result, the lifetime predicted by the series system model is pessimistic and can be improved by taking advantage of this redundancy. Our data shows that a grid can tolerate over 30 line failures, with 2-3X longer lifetimes.

On the other hand, the rate of EM degradation in power grid lines depends on the current density, and hence on the patterns of currents drawn by the underlying circuitry. It is impractical to assume that the exact current waveforms are available for all the chip workload scenarios, since this would require the simulation of the chip for millions of clock cycles at a low enough level of abstraction that would provide the current waveforms. Moreover, one might need to verify and check the grid *early* in the design flow, before fully designing the underlying circuit. Therefore, one would like a *vectorless* approach that can deal with the uncertainty about the underlying circuit currents.

In this work, we present a new EM checking approach that reduces the pessimism of SEB and which we extend to a vectorless framework for verification. Two main methods are discussed to solve the vectorless case, one is exact and theoretically interesting, and the other is approximate, fairly accurate, and much more practical.

The remainder of the paper is organized as follows. In section 2 we present a background on electromigration and the power grid model. Section 3 presents the problem definition. The proposed approaches are explained in sections 4 and 5, followed by the experimental results in section 6. Finally, section 7 concludes the paper.

2. BACKGROUND

2.1 Electromigration

Electromigration 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 can cause *depletion* of enough material so as to create an open circuit in a wire. The failure time $\mathbf{T}$ of a metal line due to electromigration is usually modeled as a random variable because degradation rates depend on the microstructure of the wire which varies due to random manufacturing variations. According to Black [2], $\mathbf{T}$ follows a *log-normal* (LN) distribution, i.e., its logarithm has a *normal* (Gaussian) distribution. The mean time-to-failure (MTF) is given by Black's equation [2, 3]:

$$\text{MTF} = \mu = \frac{wt}{A} J^{-\eta} \exp\left(\frac{E_a}{kT_m}\right)$$

*This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

where A is an experimental constant that depends on the physical properties of the metal line (volume resistivity, etc.), w and t are the width and thickness of the line respectively, 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 Boltzmann's constant, T_m 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 the conductors made of a given material.

For sufficiently short lines, the back-stress developed due to the accumulation of atoms at the ends of the line could overcome the build-up of the critical stress required for creation of a void, and thus the line is no longer susceptible to EM failure [4]. This effect, called the *Blech Effect*, is quantified in terms of a critical value β_c, where a line of length L and current density J is said to be EM-immune if $JL < \beta_c$, and EM-susceptible if $JL \geq \beta_c$.

2.2 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 under time-varying currents is to compute a constant value called the *effective-EM* current, derived from the time-varying current waveform. The value obtained represents the DC current that effectively gives the same lifetime as the original waveform under the same conditions. Power grid lines carry mostly-unidirectional currents for which, effective currents are chosen as the average currents. Accordingly, it is sufficient to consider a DC model of the grid subject to average current sources that model the currents drawn by the underlying logic blocks. This is justified because the power grid is a linear system, and hence its average branch currents can be obtained by subjecting it to average current sources.

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:

$$\mathbf{G}(t)V(t) = I$$

where $\mathbf{G}(t)$ is the conductance matrix of the grid, and $V(t)$ is the vector of voltage drops. The time dependence is introduced because $\mathbf{G}$ varies over time as grid lines start to fail due to electromigration. As long as the grid is connected, the matrix $\mathbf{G}(t)$ is known to be a diagonally-dominant symmetric positive definite $\mathcal{M}$-matrix, so that $\mathbf{G}^{-1}(t)$ exists and $\mathbf{G}^{-1}(t) \geq 0$ [5].

Because Black's model depends on the current density through the metal line, branch currents are needed. Let b be the number of branches in the grid, and let $I_{b,l}(t)$ represent the branch currents where $l \in \{1 \ldots b\}$, and let $I_b(t)$ be the vector of all branch currents. Keep in mind that the time dependence in $\mathbf{G}(t)$, $V(t)$, nd $I_b(t)$ is solely due to the change in the conductance matrix as a result of EM degradation.

Relating all the branch currents to the voltage drops across them we get:

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

where $\mathbf{R}$ is a $b \times b$ diagonal matrix of the branch resistance values and $\mathbf{M}$ 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 l^{th} branch, else a 0 occurs. The signs of the non-zero 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.

2.3 Mean Estimation by Random Sampling

We assess the reliability of the power grid by computing its MTF (or average worst-case TTF) using a *Monte Carlo* approach. Also known as *random sampling*, Monte Carlo refers to iteratively selecting specific values from the domain of a distribution, and computing their arithmetic average as an estimate of the mean.

Consider a continuous random variable $\mathbf{x}$ 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 in [6] by:

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

where s_w is the unbiased estimator of σ and $z_{\alpha/2}$ is the $(1-\alpha/2)$-percentile of the random variable $[(\bar{x}_w - \mu)/(\sigma/\sqrt{w})]$ having a *standard normal* distribution. The usage of s_n instead of σ (which is unknown) in (1), is acceptable when w is large ($w \geq 30$ as specified in [6]) because the random variable $[(\bar{x}_w - \mu)/(s_w/\sqrt{w})]$ has a t-distribution which approaches the standard normal for large w.

In the general case where the distribution is unknown (not necessarily normal), the random variable $[(\bar{x}_w - \mu)/(s_w/\sqrt{w})]$ 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. With this, one can compute the same stopping criterion in (1), which we use throughout the paper as a stopping criterion for Monte-Carlo whenever needed.

3. PROBLEM DEFINITION

Generally, it is not realistic to expect users to specify the exact 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 till 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.

The fact is, modern integrated circuits have complex *multi-modal behavior*, where major parts 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 then we can write:

$$\alpha_{\min} \leq \alpha \leq \alpha_{\max} \quad (2)$$

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:

$$I_{\ell,\min} \leq \mathbf{L}\alpha \leq I_{\ell,\max} \quad (3)$$

where $\mathbf{L}$ is a $n \times nr$ matrix such that $I = \mathbf{L}\alpha$. The matrix $\mathbf{L}$ 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 that 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 [7]. If m is the total number of global constraints, then we can write:

$$I_{g,\min} \leq \mathbf{SL}\alpha \leq I_{g,\max} \tag{4}$$

where $\mathbf{S}$ is an $m \times n$ matrix that only contains 0s and 1s and indicates which current sources are present in each global constraint.

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

$$\mathbf{B}\alpha = \mathbf{1}_n \tag{5}$$

where $\mathbf{B}$ is an $n \times nr$ matrix containing only 1s and 0s such that the vector $\mathbf{B}\alpha$ contains the sum of mode probabilities per block in each of its entries, and $\mathbf{1}_n$ is a vector of size n containing all ones. 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 (2) , (3), (4), and (5).

For example, consider a circuit having three blocks with two modes of operation each: low power and high performance. Assume that the blocks draw respectively $0.1A$, $0.2A$, and $0.15A$ on average in low power mode, and $0.2A$, $0.3A$, and $0.25A$ in high performance mode. Also, let α_{11}, α_{21}, and α_{31} denote the probabilities of the blocks being in low power mode, and α_{12}, α_{22}, and α_{32} the probabilities of being in high performance mode. The average currents I_1, I_2, and I_3 of the blocks are given as follows:

$$\begin{aligned} I_1 &= 0.1\alpha_{11} + 0.2\alpha_{12} \\ I_2 &= 0.2\alpha_{21} + 0.3\alpha_{22} \\ I_3 &= 0.15\alpha_{31} + 0.25\alpha_{32} \end{aligned}$$

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

$0.2 \leq \alpha_{11} \leq 0.6$	$0.11 \leq 0.1\alpha_{11} + 0.2\alpha_{12} \leq 0.18$
$0.1 \leq \alpha_{21} \leq 0.7$	$0.21 \leq 0.2\alpha_{21} + 0.3\alpha_{22} \leq 0.29$
$0.3 \leq \alpha_{31} \leq 0.9$	$0.17 \leq 0.15\alpha_{31} + 0.25\alpha_{32} \leq 0.24$
$0.2 \leq \alpha_{12} \leq 0.5$	$\alpha_{11} + \alpha_{12} = 1$
$0.1 \leq \alpha_{22} \leq 0.9$	$\alpha_{21} + \alpha_{22} = 1$
$0.6 \leq \alpha_{32} \leq 0.9$	$\alpha_{31} + \alpha_{32} = 1$

$$\begin{aligned} 0.35 \leq 0.1\alpha_{11} + 0.2\alpha_{12} + 0.2\alpha_{21} + 0.3\alpha_{22} \leq 0.41 \\ 0.40 \leq 0.2\alpha_{21} + 0.3\alpha_{22} + 0.15\alpha_{31} + 0.25\alpha_{32} \leq 0.48 \end{aligned}$$

The first column shows the set of constraints on individual α's. the second column shows the constraints on individual average currents (first three), and the constraints on the sums of mode probabilities (last three). The last two constraints are global constraints on $I_1 + I_2$, and $I_2 + I_3$ respectively.

For every feasible setting of α, the overall block average currents are different, and the power grid reliability (MTF) is correspondingly different. Our goal is to look for the average *worst-case* TTF of the power grid given all the possible feasible combinations of α. This raises two major questions:

1. How to estimate the MTF of the power grid for a given vector of mode probabilities $\alpha \in \mathcal{F}_\alpha$?
2. How to find the average worst-case TTF of the grid, over all the feasible vectors in $\mathcal{F}_\alpha$?

Answering these questions would lead to a framework that allows vectorless EM checking while imposing reasonable and minimal demands on the user.

4. ESTIMATING GRID FAILURE TIME

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 as weak as its weakest link. Given the mesh structure of modern power grids, it is overly pessimistic to employ the series system for EM estimation and checking. In this section, we introduce a new approach to estimate the mean time-to-failure of the grid in the theoretical case where the vector of mode probabilities α is known.

Note that, the development of the approach below, and the other material in this section (section 4) overlaps with the content of [8]. This small overlap is unavoidable due to lack of space and to make sure each paper is understandable on its own.

Generally, for a grid to function as intended, the voltage drop at each of its nodes should be smaller than a certain threshold because otherwise, soft errors in the underlying logic may occur [9]. A node is said to be *safe* when its voltage drop meets the corresponding threshold condition. Let V_{th} be the vector of all the threshold values which are typically user-specified, and assume that $V_{th} > 0$ to avoid trivial cases.

For a given α, the vector of average current sources can be found using $I = \mathbf{L}\alpha$. Assume that, for this particular I, all the nodes of the initial grid $\mathbf{G}(0)$ are safe, i.e. $V(0) \leq V_{th}$. As we move forward in time, the EM-susceptible lines start to fail in the order of their failure times due to electromigration. We model the failure of every line by an open circuit. The grid is deemed to fail at the earliest time for which the condition $V(t) \leq V_{th}$ is no longer true, meaning when any of the grid nodes becomes *unsafe*. This new model is referred to as the *mesh* model, and is used to determine the failure time of the grid when the failure times of its lines are known. Notice that modeling the failure of a line by an open circuit leads to conservative results, because typical EM models assume that a line fails when its resistance rise above a certain threshold, and hence, the line continues to conduct current after failure but with high resistance. Thus, the infinite resistance model assumes that the line is more degraded than it actually is.

In order to estimate the MTF of the power grid using the mesh model, we perform *Monte-Carlo* analysis by collecting time-to-failure samples of the power grid until the convergence condition of Monte-Carlo is met. Employing the mesh model, we need to discover the EM-immune lines in the grid, and find the MTF of the other lines. For that, current densities and hence branch currents are required. Since the grid is changing over time due to the failure of its components, the branch currents will also change. Due to lack of space, this paper will assume that the statistics of the lines can be determined using the branch currents of the grid before the failure of any of its components. The case of changing branch currents is treated in details in [8], and we are working on an extension of that work to the case of varying workloads.

If $\mathbf{G}_0$ is the conductance matrix of the original grid, i.e. $\mathbf{G}_0 = \mathbf{G}(0)$, then the vector of initial voltage drops can be written as $V_0 = V(0) = \mathbf{G}_0^{-1} I$. This allows writing:

$$I_b(0) = I_b = -\mathbf{R}^{-1}\mathbf{M}^T\mathbf{G}_0^{-1} I$$

4.1 Sampling from a Lognormal

At $t = 0$, the current density of a line l with width w_l, thickness t_l, length L_l, and branch current $I_{b,l}$, can be written as:

$$J_l = \frac{|I_{b,l}|}{w_l t_l} \tag{6}$$

To know if line l is EM-susceptible, $J_l L_l$ should be computed and compared to β_c. If $J_l L_l < \beta_c$, then the line is EM-immune and should be discarded from the analysis. Otherwise, its MTF μ_l should be computed using Black's equation which can be rewritten as follows:

$$\mu_l = \frac{(w_l t_l)^{\eta+1}}{A} |I_{b,l}|^{-\eta} \exp\left(\frac{E_a}{kT_m}\right) \tag{7}$$

For the purpose of Monte Carlo analysis, a time-to-failure (TTF) sample τ_l should be assigned to every EM-susceptible line in every Monte Carlo 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 following transformation [10]:

$$\tau_l = \mu_l \exp\left(\psi_l \sigma_{\ln} - 0.5\sigma_{\ln}^2\right) \quad (8)$$

If b_l^T is the row of $-\mathbf{R}^{-1}\mathbf{M}^T\mathbf{G}_0^{-1}$ that corresponds to line l, then $I_{b,l} = b_l^T I$ 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 (8) and (7):

$$\tau_l = |c_l^T I|^{-\eta} \quad (9)$$

where $c_l \triangleq \left[\frac{(w_l t_l)^{\eta+1}}{A} \exp\left(\frac{E_a}{kT_m}\right) \exp\left(\psi_l \sigma_{\ln} - 0.5\sigma_{\ln}^2\right)\right]^{-\frac{1}{\eta}} b_l$

4.2 Updating Voltage Drops

Checking if the grid is failing at a particular point in time requires checking the condition $V(t) \leq V_{th}$. This means that $V(t)$ should be recomputed every time a line in the grid fails. One way of doing that is by updating $\mathbf{G}$ and then resolving $V(t) = \mathbf{G}^{-1}(t)I$ using LU factorization after the failure of every line. However, this turns out to be computationally expensive. In fact, since we are modeling the failure of every line by an open circuit, we can write the change in $\mathbf{G}$ corresponding to the k^{th} line failure as a rank-1 matrix $\Delta\mathbf{G}_k$. If the k^{th} failing line has a conductance g_k and is connected between nodes x and y (with $x, y \in \{0, 1, \ldots, n\}$ and $x > y$), then $\Delta\mathbf{G}_k = -u_k u_k^T$ with:

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

where e_λ is a column vector of appropriate size containing 1 at the λ^{th} location and zeros at all other locations, with e_0 being a vector of all zeros.

Define $\mathbf{U} = [u_1 \ \ldots \ u_k]$. Clearly, $\sum_{j=1}^{k} \Delta\mathbf{G}_k = -\mathbf{U}\mathbf{U}^T$. This means that we can write the vector of voltage drops V_k after the failure of k lines as:

$$V_k = \left(\mathbf{G}_0 + \sum_{j=1}^{k} \Delta\mathbf{G}_k\right)^{-1} I = \left(\mathbf{G}_0 - \mathbf{U}\mathbf{U}^T\right)^{-1} I \quad (10)$$

The *Woodbury* formula [11] gives:

$$\left(\mathbf{G}_0 - \mathbf{U}\mathbf{U}^T\right)^{-1} = \mathbf{G}_0^{-1} + \mathbf{G}_0^{-1}\mathbf{U}(\mathbf{I}_k - \mathbf{U}^T\mathbf{G}_0^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{G}_0^{-1} \quad (11)$$

where $\mathbf{I}_k$ is the $k \times k$ identity matrix. Using (10) and (11), we have:

$$V_k = \mathbf{G}_0^{-1} I + \left[\mathbf{G}_0^{-1}\mathbf{U}(\mathbf{I}_k - \mathbf{U}^T\mathbf{G}_0^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{G}_0^{-1}\right] I$$

Define $\mathbf{Z} = \mathbf{G}_0^{-1}\mathbf{U} = [\mathbf{G}_0^{-1}u_1 \ \ldots \ \mathbf{G}_0^{-1}u_q]$. Using one LU factorization of $\mathbf{G}_0$, and $k+1$ forward/backward substitutions, $\mathbf{G}_0^{-1}I = V_0$ and $\mathbf{Z}$ can be efficiently found. Therefore,

$$V_k = V_0 + \mathbf{Z}\mathbf{W}^{-1}\mathbf{U}^T V_0$$

where

$$\mathbf{W} = \mathbf{I}_k - \mathbf{U}^T\mathbf{Z}$$

Since k is generally small, $\mathbf{W}$ is also small, which means that finding $\mathbf{W}^{-1}\mathbf{U}^T V_0$ can be efficiently done by LU factorization. Also, the matrices $\mathbf{Z}$ and $\mathbf{U}^T\mathbf{Z}$ need not be calculated from scratch for each failure, but can be updated by appending the appropriate vectors.

4.3 Singular Case

Since we are modeling the failure of a line by an open circuit, it is possible for $\mathbf{G}(t)$ to become singular as one of the nodes might become isolated. In this case the condition $V(t) \leq V_{th}$ is automatically violated because an isolated node represents a high impedance with an unknown voltage level. Therefore, the grid is deemed to fail at the earliest time for which the condition $V(t) \leq V_{th}$ is no longer true or when the conductance matrix of the grid becomes singular. It is known that $\mathbf{G}$ is singular if and only if, $\mathbf{W}$ is singular, and hence we detect the singularity of $\mathbf{G}$ by checking if $\mathbf{W}$ is invertible, which is easy since $\mathbf{W}$ is generally small. With this, we solve the case where the vector of mode probabilities α is known.

5. OPTIMIZATION

When α is unknown, the problem becomes to find the average worst-case TTF of the grid over all the feasible vectors α. For that, we start by transforming the feasible space $\mathcal{F}_\alpha$ to the current domain, which helps reduce the number of variables from nr to n, as well as the number of constraints. It is easy to see that replacing $\mathbf{L}\alpha$ by I in (3) and (4) results in the first set of constraints defining the feasible space of currents:

$$I_{\ell,\min} \leq I \leq I_{\ell,\max} \quad (12)$$

$$I_{g,\min} \leq \mathbf{S}I \leq I_{g,\max} \quad (13)$$

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 all the mode probabilities corresponding to I_j, then due to (2) 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 (2). Due to (5), 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):

$$\begin{array}{ll} \text{Min/Max} & \sum_{k=1}^{r} \alpha_{jk} I_{jk} \\ \text{subject to} & \alpha_{j,\min} \leq \alpha \leq \alpha_{j,\max} \\ & \sum_{k=1}^{r} \alpha_{jk} = 1 \end{array}$$

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 because r is small, then the LPs are very easy to solve. Ultimately if all the LPs turn out to be feasible, we get lower and upper bounds on every current source. However, (12) also provides similar bounds, hence, all the bounds should be combined to obtain:

$$I_{\min} \leq I \leq I_{\max} \quad (14)$$

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 (14) and (13). Note that $I \in \mathcal{F}$ if and only if $\alpha \in \mathcal{F}_\alpha$.

Back to the example of section 3, the resulting reduced set of constraints in the current domain would be:

$$\begin{array}{c} 0.14 \leq I_1 \leq 0.17 \\ 0.22 \leq I_2 \leq 0.25 \\ 0.21 \leq I_3 \leq 0.24 \\ 0.35 \leq I_1 + I_2 \leq 0.41 \\ 0.40 \leq I_2 + I_3 \leq 0.48 \end{array}$$

Over all the feasible vectors $I \in \mathcal{F}$, we would like to find the average worst-case TTF of the grid, which we do 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 equation (9) which in this case yields an expression for every TTF since I is not fixed.

Define Ψ to be the vector containing the samples ψ_l, $l \in \{1, \ldots, b\}$, and let $\mathcal{T}(\Psi, I)$ be a function defined on $\mathcal{F}$ such that for every vector $I \in \mathcal{F}$, $\mathcal{T}(\Psi, I)$ is the grid failure time corresponding to the set of samples in Ψ and subject to the vector of source currents I. If Ψ_i represents the vector containing the samples chosen at Monte Carlo iteration i, then the goal is to solve the following set of optimization problems:

$$\begin{array}{ll} \text{While} & \text{(Monte-Carlo has not converged) :} \\ & \text{Minimize:} \quad \mathcal{T}(\Psi_i, I) \\ & \text{subject to:} \quad I \in \mathcal{F} \end{array} \quad (15)$$

In the following, we discuss how to solve every minimization problem in the loop above given a fixed vector Ψ. We first explain the local optimization given a staring point in the feasible space, and then show both the exact and the approximate global optimizations.

5.1 Local Optimization

Let $I^{(1)}$ be a given point in $\mathcal{F}$ and $\mathcal{T}(I^{(1)})$ be the corresponding grid time-to-failure. As we will be explained later, we need several initial points in $\mathcal{F}$ to solve every iteration in (15), and every initial point will lead to a subset of $\mathcal{F}$ in which a local optimization will be performed. Therefore, a superscript is used to index the initial points that will be chosen and the corresponding subsets.

In order to compute $\mathcal{T}(I^{(1)})$, we need to compute the JL product of every line, filter out all the lines that turn out to be EM-immune, and sort the other lines according to their time-to-failure. Let $l_1^{(1)}, l_2^{(1)} \ldots, l_\zeta^{(1)}$ be the resulting sorted list of EM-susceptible lines, and $\tau_1^{(1)}, \tau_2^{(1)} \ldots, \tau_\zeta^{(1)}$ the corresponding list of TTFs such that:

$$\tau_1^{(1)} \leq \tau_2^{(1)} \leq \ldots \leq \tau_\zeta^{(1)} \tag{16}$$

Also, let $l_{\zeta+1}^{(1)}, l_{\zeta+2}^{(1)} \ldots, l_b^{(1)}$ be the list of all the other lines, i.e. the EM-immune ones.

Assume that, according to the order in (16), the grid fails for the first time after the failure of the first p EM-susceptible lines. In other words, the grid is safe if $l_1^{(1)}, l_2^{(1)}, \ldots, l_{p-1}^{(1)}$ fail, but fails when $l_p^{(1)}$ fails. This implies that $\mathcal{T}(I^{(1)}) = \tau_p(I^{(1)})$ as explained in section 4. Throughout the rest of the paper, we assume that $p < \zeta$, because otherwise, the grid becomes immortal indefinitely which is unrealistic. For convenience, we omit the superscript notation in this section and refer to $l_1^{(1)}, \ldots, l_b^{(1)}$ by $l_1, \ldots, l_b$, and to $\tau_1^{(1)}, \ldots, \tau_\zeta^{(1)}$ by $\tau_1, \ldots, \tau_\zeta$.

5.1.1 General Case

Define $\mathbf{G}_p$ to be the conductance matrix of the grid after the failure of $l_1, \ldots, l_p$, and $\mathbf{G}_{p-1}$ the conductance matrix after the failure of $l_1, \ldots, l_{p-1}$ only. In this section, we assume that the failure occurs due to the violation of the voltage drop condition. The case of singularity is discussed later.

For $I \in \mathcal{F}$, assume that $\tau_i(I)$ can be written using (9) as:

$$\tau_i(I) = |c_i^T I|^{-\eta}$$

Let $\xi_i = \pm 1$ denote the sign of $c_i^T I^{(1)}$ for $i \in \{1, \ldots, p\}$, i.e. $\xi_i = \frac{c_i^T I^{(1)}}{|c_i^T I^{(1)}|}$, which implies $\xi_i c_i^T I^{(1)} \geq 0$.

Definition 1. *We define $\mathcal{S}^{(1)}$ to be the subset of $\mathcal{F}$ corresponding to $I^{(1)}$ such that for every $I \in \mathcal{S}^{(1)}$:*

1. *Lines $l_1, \ldots, l_p$ are EM-susceptible while lines $l_{\zeta+1}, \ldots, l_b$ are EM-immune.*
2. *The branch currents of $l_1, \ldots, l_p$ have the same direction they have for $I = I^{(1)}$*
3. *Line l_p fails right after the set of lines $\{l_1, \ldots, l_{p-1}\}$ and before all the other lines.*
4. *Line l_p is the first line to cause the failure of the grid.*

It is easy to check that the point $I^{(1)}$ belongs to $\mathcal{S}^{(1)}$ as all of the four conditions are clearly satisfied.

Claim 1. *For every $I \in \mathcal{S}^{(1)}$, $\mathcal{T}(I) = \tau_p(I)$*

Proof. For any $I \in \mathcal{S}^{(1)}$, l_p fails right after the failure of the lines in the set the set $\{l_1, \ldots, l_{p-1}\}$ and is the first to cause the failure of the grid. Therefore, $\mathcal{T}(I) = \tau_p(I)$. □

Claim 1 shows how to write $\mathcal{T}(I)$ as a closed form expression inside $\mathcal{S}^{(1)}$ which is a non-empty subset because $I^{(1)} \in \mathcal{S}^{(1)}$. To minimize $\mathcal{T}$ in $\mathcal{F}$, we start by performing a *local* optimization in a subset of $\mathcal{F}$ where $\mathcal{T}$ can be explicitly defined, namely $\mathcal{S}^{(1)}$. In other words, we are interested in solving the following optimization problem:

$$\begin{array}{ll} \text{Minimize} & \tau_p(I) \\ \text{subject to} & I \in \mathcal{S}^{(1)} \end{array} \tag{17}$$

To solve (17), we introduce a new function ν_p on $\mathcal{S}^{(1)}$ defined by $\nu_p(I) = (\tau_p(I))^{-\frac{1}{\eta}}$. Since $\forall I \in \mathcal{S}^{(1)}, \xi_p c_p^T I \geq 0$, we can write $\nu_p(I) = (\tau_p(I))^{-\frac{1}{\eta}} = \left[|c_p^T I|^{-\eta}\right]^{-\frac{1}{\eta}} = |c_p^T I| = \xi_p c_p^T I$, meaning, $\nu_p(I)$ is a linear function of I.

Lemma 1. *A point I_{opt} is a solution for (17) if, and only if, $\nu_p(I_{opt})$ is a maximum for $\nu_p(I)$ on $\mathcal{S}^{(1)}$.*

Proof. Since $\eta > 0$ and $\tau_p(I)$ is a positive function for $I \in \mathcal{S}^{(1)}$ (meaning $\nu(I)$ is a positive function in $\mathcal{S}^{(1)}$ as well), then:

$$\begin{aligned} &I_{\text{opt}} \text{ is a solution for (17)} \\ &\Leftrightarrow \tau_p(I) \geq \tau_p(I_{\text{opt}}), \forall I \in \mathcal{S}^{(1)} \\ &\Leftrightarrow (\nu_p(I))^{-\eta} \geq (\nu_p(I_{\text{opt}}))^{-\eta}, \forall I \in \mathcal{S}^{(1)} \\ &\Leftrightarrow \left[(\nu_p(I))^{-\eta}\right]^{-\frac{1}{\eta}} \leq \left[(\nu_p(I_{\text{opt}}))^{-\eta}\right]^{-\frac{1}{\eta}}, \forall I \in \mathcal{S}^{(1)} \\ &\Leftrightarrow \nu_p(I) \leq \nu_p(I_{\text{opt}}), \forall I \in \mathcal{S}^{(1)} \end{aligned}$$

which proves the lemma. □

Lemma 1 implies that in order to solve (17), we can solve instead the following maximization problem to get I_{opt}:

$$\begin{array}{ll} \text{Maximize} & \xi_p c_p^T I \\ \text{subject to} & I \in \mathcal{S}^{(1)} \end{array} \tag{18}$$

The solution to (17) would simply be $(\xi_p c_p^T I_{\text{opt}})^{-\eta}$.

For $i \in \{1, \ldots, b\}$, let b_i denotes the row of $-\mathbf{R}^{-1}\mathbf{M}^T\mathbf{G}_0^{-1}$ that corresponds to line l_i, and let w_i, t_i, and L_i denote the width, thickness, and length of l_i respectively.

Theorem 1. *$I \in \mathcal{S}^{(1)}$ if, and only if, the following constraints are satisfied:*

$$\frac{L_i}{w_i t_i} \xi_i b_i^T I \geq \beta_c \qquad \text{for } i \in \{1, \ldots, p\} \tag{19}$$

$$\frac{L_i}{w_i t_i} |b_i^T I| < \beta_c \qquad \text{for } i \in \{\zeta + 1, \ldots, b\} \tag{20}$$

$$\xi_p c_p^T I - \xi_i c_i^T I \leq 0 \qquad \text{for } i \in \{1, \ldots, p-1\} \tag{21}$$

$$|c_i^T I| - \xi_p c_p^T I \leq 0 \qquad \text{for } i \in \{p+1, \ldots, \zeta\} \tag{22}$$

$$\mathbf{G}_{p-1}^{-1} I \leq V_{th} \tag{23}$$

$$\mathbf{G}_p^{-1} I \not\leq V_{th} \tag{24}$$

The proof theorem 1 can be found in the appendix.

The inequalities in (19), (21), (22), and (23) can be combined in matrix form as $\mathbf{H}_1 I \leq h_1$, and the inequalities in (20) can be combined as $\mathbf{H}_2 I < h_2$, such that the number of rows in $\mathbf{H}_1$ is $\gamma_1 = n + 2\zeta - 1$, and the number of rows in $\mathbf{H}_2$ is $\gamma_2 = 2(b - \zeta)$. The details are skipped due to lack of space. Ultimately, $\mathcal{S}^{(1)}$ can be redefined as follows:

$$\begin{array}{rcl} \mathbf{H}_1 I & \leq & h_1 \\ \mathbf{H}_2 I & < & h_2 \\ \mathbf{G}_p^{-1} I & \not\leq & V_{th} \end{array}$$

All the inequalities presented above are linear and define a convex polytope (or the interior of a convex polytope) except $\mathbf{G}_p^{-1} I \not\leq V_{th}$ which consists of a *disjunction of constraints* where at least one entry of $\mathbf{G}_p^{-1} I$ has to be greater than its corresponding entry in V_{th}. We deal with that using a theorem that will be presented shortly.

Throughout the rest of the paper, we will be using the 1-norm and the infinity norm defined as follows: given a vector $x \in \mathbb{R}^n$ with entries x_i, $i \in \{1, \ldots, n\}$:

$$\|x\|_1 \triangleq \sum_{i=1}^n |x_i| \qquad \|x\|_\infty \triangleq \max_{i=1\ldots n} |x_i|$$

For any strictly positive number δ define $d = \|V_{th}\|_\infty (1 + \delta)$.

Algorithm 1 EXACT_GLOBAL_MINIMIZATION

Input: Ψ
Output: Global Minimum of $\mathcal{T}(\Psi, I)$
1: Find $I^{(1)}$ in $\mathcal{F}$ using an LP
2: Set $solved \leftarrow false$ and $i \leftarrow 2$
3: **while** $(solved = false)$ **do**
4: Find $\mathcal{T}(\Psi, I^{(i-1)})$
5: Build the constraints defining $\mathcal{S}^{(i-1)}$ as outline in section 5.1
6: Solve $\max_{I \in \mathcal{S}^{(i-1)}} \mathcal{T}(I)$ using an ILP
7: Solve a feasibility problem in $\mathcal{F} - \cup_{j=1}^{i-1} \mathcal{S}^{(j)}$ to get $I^{(i)}$ as outlined in section 5.2
8: **if** ($I^{(i)}$ cannot be found) **then**
9: $solved \leftarrow true$
10: **end if**
11: $i \leftarrow i + 1$
12: **end while**
13: Return the smallest local minimum found

Theorem 2. *$\mathbf{G}_p^{-1} I \not\leq V_{th}$ if, and only if, $\exists y \in \{0,1\}^n$ with $\|y\|_1 \leq n-1$ such that $\mathbf{G}_p^{-1} I > V_{th} - dy$.*

Proof. If $\mathbf{G}_p^{-1} I \not\leq V_{th}$, then there exists a non-empty set of indices $\mathcal{K} \subseteq \{1, \ldots, n\}$ such that $e_k^T \mathbf{G}_p^{-1} I > e_k^T V_{th}$ for every $k \in \mathcal{K}$. If we let $y = [y_1 \ \ldots \ y_n]^T$ with $y_k = 0$ for $k \in \mathcal{K}$, and $y_k = 1$ otherwise, then clearly $\|y\|_1 \leq n-1$, and $\mathbf{G}_p^{-1} I > V_{th} - dy$ because $d > \|V_{th}\|_\infty$ and the entries of the vector $\mathbf{G}_p^{-1} I$ are always positive.

On the other hand, if there exists $y \in \{0,1\}^n$ with $\|y\|_1 \leq n-1$, and $\mathbf{G}_p^{-1} I > V_{th} - dy$, then $\exists k$ such that $y_k = 0$, and $e_k^T \mathbf{G}_p^{-1} I > e_k^T V_{th}$. Therefore, $\mathbf{G}_p^{-1} I \not\leq V_{th}$. □

The theorem above allows replacing the last constraint in theorem 1 by

$$\begin{array}{rcl} \mathbf{G}_p^{-1} I & > & V_{th} - dy \\ \|y\|_1 & \leq & n-1 \\ y & \in & \{0,1\}^n \end{array} \tag{25}$$

This shows that (18) is an integer linear program (ILP) because it has a linear objective function, linear constraints, and some integer variables, namely the entries of y. Solving (18) would solve (17) as explained before, and would minimize $\mathcal{T}$ inside $\mathcal{S}^{(1)} \subset \mathcal{F}$.

5.1.2 Singular Case

If at $I^{(1)}$ the grid fails by singularity, then the same analysis as above can be done. The only difference is that the constraint $\mathbf{G}_p^{-1} I \not\leq V_{th}$ cannot be added, and is in fact redundant because $\mathbf{G}_p$ is known to be singular in this case, and there is no need to add that as one of the constraints defining $\mathcal{S}^{(1)}$. Notice that in this case, (18) becomes a linear program.

In the following, we show how to globally minimize $\mathcal{T}$ in $\mathcal{F}$ by performing a set of local optimizations as above until $\mathcal{F}$ is fully explored.

5.2 Exact Global Optimization

Similarly to $\mathcal{S}^{(1)}$, every local optimization requires a starting point in $\mathcal{F}$. In order to create $\mathcal{S}^{(2)}$, we need a point $I^{(2)} \in \mathcal{F}$ at which we compute the TTF of the grid and then follow a similar procedure to the one explained in the previous section. Note that finding $I^{(1)}$ can be done by solving a linear feasibility problem in the set $\mathcal{F}$. However, we cannot do the same for $I^{(2)}$ because if we choose $I^{(2)}$ in $\mathcal{F}$ without other restrictions and it turns out that it belongs to $\mathcal{S}^{(1)}$, then $\mathcal{S}^{(2)}$ becomes identical to $\mathcal{S}^{(1)}$, which adds redundancy to our approach and accordingly, the global optimization may or may not terminate. In short, $I^{(2)}$ should be chosen in the set $\mathcal{F} - \mathcal{S}^{(1)}$. Using the constraints in theorem 1, we can infer the set of conditions required for $I^{(2)}$ to be in outside $\mathcal{S}^{(1)}$ (in the general case) as follows:

$$\begin{array}{rcl} \mathbf{H}_1 I & \not\leq & h_1 \\ \text{or} \quad \mathbf{H}_2 I & \not< & h_2 \\ \text{or} \quad \mathbf{G}_p^{-1} I & \leq & V_{th} \end{array} \tag{26}$$

For any strictly positive number δ, define

$$d = (1+\delta) \max\left(\|h_1 - \mathbf{H}_1 I\|_\infty, \|h_2 - \mathbf{H}_2 I\|_\infty, \|\mathbf{G}_p^{-1} I - V_{th}\|_\infty \right)$$

Theorem 3. *(26) is true if and only if, there exists $x_1 \in \{0,1\}^{\gamma_1}$, $x_2 \in \{0,1\}^{\gamma_2}$, and $x_3 \in \{0,1\}$ with*

$$\|x_1\|_1 + \|x_2\|_1 + x_3 \leq \gamma_1 + \gamma_2$$

such that:

$$\begin{array}{rcl} \mathbf{H}_1 I & > & h_1 - dx_1 \\ \mathbf{H}_2 I & \geq & h_2 - dx_2 \\ \mathbf{G}_p^{-1} I & \leq & V_{th} + dx_3 \mathbf{1}_n \end{array} \tag{27}$$

The proof is skipped due to lack of space, and is in fact very similar to that of theorem 2.

In the singular case, the last set of constraints in (27) as well as the binary variable z are not needed. Theorem 3 implies that finding $I^{(2)}$ requires solving a feasibility problem in the space defined by (27), which can be done using an ILP. The same approach should be used to find the i^{th} starting point corresponding to subset $\mathcal{S}^{(j)}$: $I^{(i)}$ should be chosen as to satisfy the constraints $I \in \mathcal{F}$ and $I \notin \mathcal{S}^{(j)}$, $j \in \{1, \ldots, i-1\}$. This can also be done using an ILP similarly to $I^{(2)}$. When such point cannot be found, i.e. when $\mathcal{F} - \cup_{j=1}^{i-1} \mathcal{S}^{(j)}$ becomes empty, we infer that the feasible space $\mathcal{F}$ is fully explored and the algorithm terminates while returning the best local minimum found. The result is one sample TTF for the grid which should be added to the other samples found in other Monte Carlo iterations. Algorithm 1 shows how to compute exact global minimum of the grid TTF given a set of normal samples Ψ using the proposed approach.

Algorithm 2 APPROXIMATE_GLOBAL_MINIMIZATION

Input: Ψ
Output: Global Minimum of $\mathcal{T}(\Psi, I)$
1: Find a starting point I_0 in $\mathcal{F}$ using an LP
2: Compute $\mathcal{T}(\Psi, I_0)$
3: Set $k \leftarrow 0$ and choose and initial temperature T_0
4: **while** $(T_{k+1} \geq T_\epsilon)$ **do**
5: Sample a new point I'_{k+1} in $\mathcal{F}$ as explained in section 5.3.1
6: Find $\mathcal{T}(\Psi, I'_{k+1})$
7: Find I_{k+1} based on the acceptance function as in (28)
8: Find T_{k+1} using (29)
9: Set $k \leftarrow k+1$
10: **end while**
11: Solve a local minimization at the best point found using a convex relaxation of the ILP and return the result (procedure of section 5.1).

5.3 Simulated Annealing with ILP Relaxation

We now present an approximate solution to the global optimization problem based on *Simulated Annealing* (SA) [12]. SA employs a neighborhood search that occasionally allows uphill moves in order to avoid being stuck at a local minimum. It 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*. We propose using both SA and the local optimizer developed before to obtain a good approximation of $\min_{I \in \mathcal{F}} \mathcal{T}(I)$.

As our acceptance function, we use the Metropolis function:

$$\mathcal{A}(I_k, I'_{k+1}, T_k) = \min\left\{1, \exp\left(-\frac{\mathcal{T}(I'_{k+1}) - \mathcal{T}(I_k)}{T_k}\right)\right\} \tag{28}$$

Table 1: Speed and Accuracy Comparisons

Power Grid		Exact Solution		Simulated Annealing		Error
Name	Nodes	Avg Min TTF (yrs)	CPU Time	Avg Min TTF (yrs)	CPU Time	
G1	141	7.50	3.21 min	7.39	2.12 min	-1.47%
G2	177	9.11	5.94 min	9.54	2.13 min	4.72%
G3	237	10.38	25.10 min	9.91	2.44 min	-4.53%
G4	392	8.01	49.24 min	7.93	2.69 min	-1.00%
G5	586	9.79	4.42 hrs	10.05	2.58 min	2.66%
G6	2,213	-	-	9.20	2.19 min	-
G7	8,413	-	-	11.79	19.34 min	-
G9	18,678	-	-	8.40	49.93 min	-
G10	32,554	-	-	7.17	1.40 hrs	-
G11	50,444	-	-	9.22	2.54 hrs	-
G12	72,692	-	-	6.40	4.83 hrs	-
G14	131,294	-	-	7.07	13.17 hrs	-

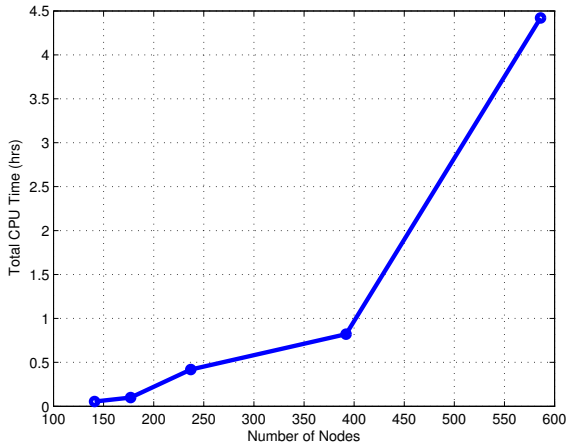

Figure 1: CPU Time of the Exact Approach versus the Number of Grid Nodes

where I_k is the current point, I'_{k+1} is the candidate point, and t_k is the current temperature. Notice that, if $\mathcal{T}(I'_{k+1}) \leq \mathcal{T}(I_k)$, then the acceptance function returns 1, and hence the new current point is $I_{k+1} = I'_{k+1}$ (accepted with a probability equal to 1). Otherwise, I'_{k+1} is accepted with a probability that depends on T_k and the difference between $\mathcal{T}(I_k)$ and $\mathcal{T}(I'_{k+1})$; in this case, accepting I'_{k+1} is referred to as *hill climbing*. For our cooling schedule, we use

$$T_k = \mathcal{K}(T_0, k) = a^{\lfloor \frac{k}{M} \rfloor} T_0 \tag{29}$$

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.

5.3.1 Sampling Random Points

Sampling a starting point I_0 in $\mathcal{F}$ can be done using a linear program. Finding the next candidate point I'_{k+1} can be done by generating a random direction θ in space such that $\|\theta\|_2 = 1$, and then computing the set Λ defined as follows:

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

Because $\mathcal{F}$ is a convex polytope, it can be shown that, if $I_k \in \mathcal{F}$, then the set Λ can be written as

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

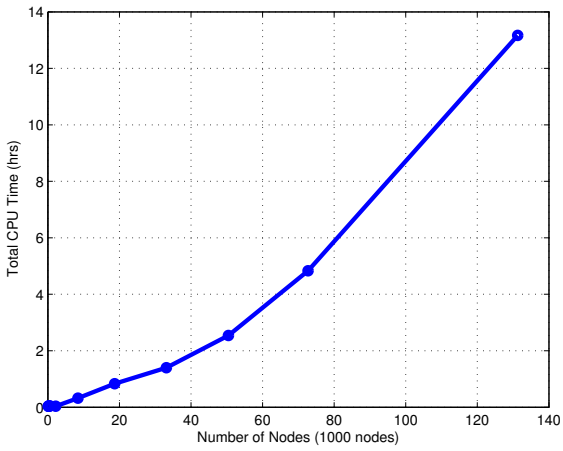

Figure 2: CPU Time of Simulated Annealing Approach versus the Number of Grid Nodes

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 $I'_{k+1} = I_k + \lambda\theta$.

5.3.2 Convex Relaxation of the ILP

Once SA terminates, we run the local optimization procedure explained in section 5.1 with the best point found by SA as a starting point and while solving an LP relaxation of (17). This can be done by allowing the entries of the vector y in (25) to be in the range $[0, 1]$ instead of the set $\{0, 1\}$. The result is an estimate sample for the minimum grid TTF. As before, enough samples should be collected until Monte Carlo converges. Algorithm 2 shows how to compute the global minimum of the grid TTF given a set of normal samples Ψ using SA and the local minimizer.

6. EXPERIMENTAL RESULTS

Algorithms 1 and 2 have been implemented in `C++`. The algorithms use the Mosek optimization package [13] to solve the required LPs and ILPs. An approximate sparse inverse of $\mathbf{G}_0$ is found using SPAI [14], and all the other required inverses are found using the Woodbury formula, i.e. (11). We carried out several experiments using 14 different power grids generated as per user specifications, including grid dimensions, metal layers, pitch and width per layer. Supply voltages and current sources were randomly placed on the grid which is assumed to have Aluminum interconnects. The parameters of the grids are consistent with 1.1V 65nm CMOS technology. As for the EM model employed, and because Aluminum is assumed, we use an activation energy of $0.9eV$, a current exponent $\eta = 1$, a nominal temperature $T_m = 373K$, a critical Blech product $\beta_c = 3000A/cm$. The lognormal standard deviation we use is $\sigma_{\ln} = 0.3$ as in [3]. All the experiments were carried out on a 2.6GHz Linux machine with 24GB of RAM. To assess the quality of our results, we computed the average worst-case grid TTF using both the exact and the SA approaches together with the required CPU time for every grid. Table 1 shows the speed and accuracy of the simulated annealing approach where the last column shows the percentage error between the mean found using each method. We can see that the error is always less than $\pm 5\%$ while the run time is much less for the SA approach. The exact approach requires solving several ILPs and for that reason the required CPU time is relatively large even for very small grids as shown in figure 1. On the other hand, the SA approach seems fairly accurate and very scalable since the run time is slightly super-linear as shown in figure 2. In fact, the *observed empirical* complexity of the implemented algorithm is found to be around $\mathcal{O}(n^{1.4})$.

To show the importance of our approach, we computed the se-

ries system TTF at the optimal points obtained, and we observed that the mesh model TTF is 2-3X larger than that of the series system. In addition, we computed the average best-case TTF using Simulated Annealing, and we noticed that there is a large separation between the average minimum and average maximum TTF for $I \in \mathcal{F}$. For instance, the separation was found to be 6.55 years for G10, 8.33 years for G11, 4.45 years for G12, and 3.8 years for G12. The large separation means that the MTF of the grid is highly sensitive to the change in currents, i.e. it is not enough to compute the MTF of the grid at an arbitrary feasible point; computing the minimum is necessary.

One might argue that the grids above are small when compared to full-chip grids containing millions of nodes. However, our approach is important for at least three reasons: 1) it is an approach that checks for EM safety using a less pessimistic model in a truly vectorless approach, 2) our method 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, and 3) almost 30% of the run time in the case of SA is taken by SPAI to compute $\mathbf{G}_0^{-1}$ which, together with the other parts of the algorithm, can be easily parallelized due the inherent independence in both SPAI and Monte Carlo iterations.

7. CONCLUSION

We described an early vectorless approach for power grid electromigration checking under a constraint-based framework to capture workload uncertainty. We presented an exact, theoretically interesting approach, and another approximate, fairly accurate, and practical approach based on Simulated Annealing. The exact approach requires solving several ILPs and turns out to be very expensive, while the SA based approach requires few LPs and one linear system solve. For small grids, the error incurred by the SA approach was relatively very small, and never exceeding $\pm 5\%$.

8. REFERENCES

[1] J. Kitchin. Statistical electromigration budgeting for reliable design and verification in a 300-MHz microprocessor. In *Symposium on VLSI Circuits Digest*, pages 115–116, 1995.

[2] J. R. Black. Electromigration, a brief survey and some recent results. *IEEE transactions on Electronic Devices*, 1969.

[3] A. Christou and M. Peckerar. *Electromigration and Electronic Device Degradation*. John Wiley and Sons, 1994.

[4] I. A. Blech. Electromigration in thin aluminium on titanium nitride. *Journal of Applied Physics*, 47(4):1203–1208, 1976.

[5] F. N. Najm. *Circuit Simulation*. John Wiley and Sons, 2010.

[6] J. E. Freund, I. R. Miller, and R. Johnson. *Probability and Statistics for Engineers*. Prentice-Hall, sixth edition, 2010.

[7] D. Kouroussis and F. N. Najm. A static patter-independent technique for power grid voltage integrity verification. In *ACM/IEEE Design Automation Conference*, Anaheim, CA, June 2003.

[8] S. Chatterjee, M. Fawaz, and F. N. Najm. Redundancy-aware electromigration checking for mesh power grids. In *IEEE/ACM International Conference On Computer-Aided Design*, San Jose, CA, Nov. 18-21 2013.

[9] 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, Nov. 2003.

[10] E. A. Amerasekera and F. N. Najm. *Failure Mechanisms in Semiconductor Devices*. John Wiley and Sons, second edition, 1997.

[11] N. J. Higham. *Functions of Matrices: theory and computation*. SIAM, first edition, 2008.

[12] M. Locatelli. *Handbook of Global Optimization*, volume 2, chapter Simulated Annealing Algorithms for Continuous Global Optimization, pages 179–229. Kluwer Academic Publishers, 2002.

[13] The MOSEK optimization software (www.mosek.com).

[14] N. H. Abdul Ghani and F. N. Najm. Fast vectorless power grid verification using an approximate inverse technique. In *ACM/IEEE Design Automation Conference*, San Fransisco, CA, Jul. 26-31 2009.

APPENDIX

A. PROOF OF THEOREM 1

Using definition 1, we know that for every $I \in \mathcal{S}^{(1)}$, lines $l_1, \ldots, l_p$ have the same direction they have for $I = I^{(1)}$, therefore:

$$\xi_i b_i^T I \geq 0 \text{ and } \xi_i c_i^T I \geq 0 \text{ for } j \in \{1, \ldots, p\} \tag{30}$$

Also, lines $l_1, \ldots, l_p$ are EM-susceptible, meaning their JL products are greater than β_c. Using (6), this can be written as:

$$\frac{|b_i^T I|}{w_i t_i} L_i \geq \beta_c, \ i = 1, \ldots, p \tag{31}$$

Similarly, lines $l_{\zeta+1}, \ldots, l_b$ are EM-immune, meaning we can write:

$$\frac{|b_i^T I|}{w_i t_i} L_i < \beta_c, \ i = \zeta + 1, \ldots, b$$

which is identical to (20).
Moreover, l_p fails after lines $\{l_1, \ldots, l_{p-1}\}$, and before all the other lines. This is equivalent to:

$$\max_{j \in \{1,\ldots,p-1\}} \tau_j(I) \leq \tau_p(I) \leq \min_{k \in \{p+1,\ldots,\zeta\}} \tau_k(I)$$

which is also equivalent to:

$$\max_{j \in \{1,\ldots,p-1\}} |c_j^T I|^{-\eta} \leq |c_p^T I|^{-\eta} \leq \min_{k \in \{p+1,\ldots,\zeta\}} |c_k^T I|^{-\eta} \tag{32}$$

The grid is safe after the failure of the lines in the set $\{l_1, \ldots, l_{p-1}\}$. This is equivalent to (23) because $\mathbf{G}_{p-1}$ is defined to be the conductance matrix of the grid after the failure of those lines.
The grid fails after the failure of the lines in the set $\{l_1, \ldots, l_p\}$. This is equivalent to (24) because $\mathbf{G}_p$ is defined to be the conductance matrix of the grid after the failure of those lines.
It remains to show that:

$$(30), (31), \text{ and } (32) \Leftrightarrow (19), (21), \text{ and } (22)$$

We do this using a two way proof. Assume (30), (31), and (32) are true, then (31) implies (19) because $|b_i^T I| = \xi_i b_i^T I$ for $i \in \{1, \ldots, p\}$. Also, (32) implies:

$$\max_{j \in \{1,\ldots,p-1\}} \left(\xi_j c_j^T I\right)^{-\eta} \leq \left(\xi_p c_p^T I\right)^{-\eta} \leq \min_{k \in \{p+1,\ldots,\zeta\}} |c_k^T I|^{-\eta} \tag{33}$$

By taking the $\left(-\frac{1}{\eta}\right)^{th}$ power of all the terms of (33), we can write

$$\max_{k \in \{p+1,\ldots,\zeta\}} |c_k^T I| \leq \xi_p c_p^T I \leq \min_{j \in \{1,\ldots,p-1\}} \xi_j c_j^T I \tag{34}$$

which implies (21) and (22). This is true because $-\frac{1}{\eta} < 0$ and hence, taking the $\left(-\frac{1}{\eta}\right)^{th}$ power reverses all the inequalities in which case, the min operator becomes a max, and vice versa.

One the other hand, assume (19), (21), and (22) are true, then (34) is true. Since

$$\max_{k \in \{p+1,\ldots,\zeta\}} |c_k^T I| \geq 0$$

then

$$0 \leq \xi_p c_p^T I \leq \min_{j \in \{1,\ldots,p-1\}} \xi_j c_j^T I$$

and hence, (30) is true. In addition, we can take the $(-\eta)^{th}$ power of all the terms in (34) to get (33) because $-\eta < 0$. Now, we can easily get (32) from (33) because (30) is true. Finally, (31) is also true because of (30) and (19). $\square$