Fast Physics-Based Electromigration Checking for On-Die Power Grids *

Sandeep Chatterjee[†], Valeriy Sukharev[$] and Farid N. Najm[†]
[†]ECE Department, University of Toronto, Toronto, ON, Canada, M5S 3G4
[$]Mentor Graphics Corporation, Fremont, CA 94538, USA
sandeep.chatterjee@mail.utoronto.ca, Valeriy_Sukharev@mentor.com,
f.najm@utoronto.ca

ABSTRACT

Due to technology scaling, electromigration (EM) signoff has become increasingly difficult, mainly due to the use of inaccurate methods for EM assessment, such as the empirical Black's model. In this paper, we present a novel approach for EM checking using physics-based models of EM degradation, which effectively removes the inaccuracy, with negligible impact on run-time. Our main contribution is to extend the existing physical models for EM in metal branches to track the degradation in multi-branch interconnect trees. We also propose effective filtering and predictor-based schemes to speed up our implementation, with minimal impact on accuracy. Our results, for a number of IBM power grid benchmarks, confirm that Black's model is overly inaccurate. The lifetimes found using our physics-based approach are on average 3x longer than those based on a (calibrated) Black's model, such as currently used in industry. For the two largest IBM benchmarks (700K branches each), our runtime is comparable to that of the Black's based approach, requiring 3 hours for the largest grid.

Keywords

Power Grids, Electromigration, hydrostatic stress, Reliability.

1. INTRODUCTION

As a result of continued scaling of integrated circuits technology, electromigration (EM) has become a major reliability concern for the design of on-die power grids in large integrated circuits. Standard practice in the industry is to break up a general grid metal structure, typically a so-called *interconnect tree* (defined later) into branches, to assess the reliability of each branch *separately* using Black's model [2] and then use the earliest branch failure time as the failure time for the whole grid. This approach is highly *inaccurate*, for many reasons. First, it ignores material flow throughout the tree, from branch to branch. As a result, if the individual branches happen to be short so that they are deemed immortal due to the Blech effect in short lines, then the tree would appear to be immortal, which is highly *optimistic* and can be entirely misleading for design. In fact, due to material flow across the tree, failures can and do happen even if the branches are short. On the other hand, because the assumption of no material flow between branches effectively means that the reliability of nearby metal lines are independent of each other, then the traditional approach can also be highly *pessimistic*, as we will see in this paper. Indeed, two identical *connected* lines that carry the same current density can in practice have quite different values of mean time-to-failure (MTF), as Gan et. al. [5] have found, so that connected lines can in fact influence each other leading to different failure times. Due to the resulting pessimism, designers are faced with an ever-shrinking margin between the predicted EM stress and the stress allowed by EM design rules. This makes EM sign-off very hard to achieve and causes overuse of metal resources in the grid.

Furthermore, traditional EM checking assumes that a power grid fails as soon as any one of its lines fails. This has been referred to as a *series model* of grid failure, and its main problem is that it ignores the inherent redundancy in the many parallel paths of the power grid. As an alternative, the *mesh model* [3] has been proposed to account for this redundancy, and its key feature is that a grid is deemed to have failed, not when the first line fails, but when the voltage drop at any grid node exceeds a user specification. However, [3] still used the Black's model to compute the EM degradation of individual lines. Thus, in order to get a more realistic estimate of grid reliability, one needs to use a mesh model *and* abandon Black's model in favor of more physical EM models.

Indeed, some approaches based on physical EM models have been recently proposed. Huang et. al. [7] have proposed an adaptation of Korhonen's physical EM models to interconnect trees. But unfortunately, the resulting approach is slow, requiring up to 32 hours to estimate the failure time of a 400K node grid. In the approach by Li et. al. [10], the time to failure of a junction (where multiple branches meet in a tree) is found by replacing its connected branches with semi-infinite limbs. Thus, atomic flow across the whole tree is not accurately accounted for.

In this paper, we propose a fast physics-based EM checking approach that accounts for material flow and coupling of stress in interconnect trees, allowing arbitrary complex geometries. We begin with the one-dimensional (1D) physical model for EM degradation within branches proposed by Korhonen [8], and extend it by introducing boundary laws at junctions so that it can track material flow and stress evolution in multi-branch interconnect trees. We refer to this as the *extended Korhonen's model*. The extended model starts out as a system of partial differential equations (PDE) coupled by the boundary laws, which we then scale and discretize to reduce it to a system of ordinary differential equations (ODE). We numerically solve this ODE system at successive time-points to track the stress evolution and find the corresponding time of void nucleation(s). Even though we're mainly interested in the mesh model approach, we report both the series model based failure time, as the earliest nucleation time in any tree, and the mesh model based failure time, as the earliest time when a node voltage drop violation is observed somewhere.

The random nature of EM degradation is accounted for by using a Monte Carlo method, in which successive samples of grid time-to-failure are found, until the estimate of the overall MTF has converged. We speed up our computation by using a filtering scheme that estimates up-front the set of trees that are most likely to impact the MTF assessment of the grid, a scheme which we will show has minimal impact on accuracy. Further, we also intro-

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

ICCAD '16, November 07-10, 2016, Austin, TX, USA
© 2016 ACM. ISBN 978-1-4503-4466-1/16/11...$15.00
DOI: http://dx.doi.org/10.1145/2966986.2967041

duce a predictive scheme that allows for faster MTF estimation by extrapolating the solution (stress curve) obtained from a few initial time-points. Testing this approach on the IBM grid benchmarks, with the largest grid up to $700K$ nodes, shows that the MTFs estimated using our physics-based approach are on average 3x longer than those based on a (calibrated) Black's model. This justifies the claim that Black's model can be overly inaccurate for modern power grids and confirms the need for physical models. With a run-time of 3 hours for the largest grid (700K nodes), this approach appears to be suitable for large VLSI circuits.

The remainder of the paper is organized as follows. In Section 2, we present some relevant background material regarding electromigration, numerical methods for solving PDEs and statistical methods which we will use later in the paper. Section 3 develops the extended Korhonen's model and Section 4 details the numerical technique that we use for solving the extended model. Section 5 describes our overall power grid EM checking approach. Section 6 outlines the implementation details and presents the experimental results. Finally, Section 7 concludes the paper.

2. BACKGROUND

2.1 Electromigration basics

Electromigration is the mass transport of metal atoms due to momentum transfer between electrons (driven by an electric field) and the atoms in a metal line. The process of EM degradation can be divided into two phases: *void nucleation* and *void growth*.

Under conditions of high current density, metal atoms are pushed in the direction of the electronic current. This creates tensile stress at the cathode and compressive stress at the anode (recall, electric current in the line flows from anode to cathode). The amount of compressive stress needed to cause an extrusion (and eventual short circuit to a neighboring line) is very high in modern metal systems, hence lines don't usually fail due to short circuits. However, the tensile stress building up at the cathode eventually leads to void nucleation when the stress reaches a critical threshold. This phase of EM degradation, when stress is increasing over time but the void has not yet nucleated, is called as the *void nucleation* phase. In this phase, the resistance of a line remains roughly the same as that of a fresh (undamaged) line.

Once a void nucleates at the cathode, the *void growth* phase begins. In some cases, depending on the geometry of the line, void nucleation by itself might be enough to cause line failure due to open circuit [9]. In other cases, a line may still continue to conduct current through the barrier material surrounding the metal, thus it is not open circuited. With time, the void starts to grow in the direction of the electronic current and the line resistance increases towards some steady-state value. Line failure is deemed to happen when the increase in line resistance reaches a critical threshold, usually set at $10\% - 20\%$ of the initial value.

2.2 The Korhonen Model

Korhonen et al. [8] proposed a one-dimensional (1D) model to describe the mechanical stress arising under the influence of electromigration. Consider a uniform metal line embedded in a rigid dielectric, that carries a current density $j(t)$. We are interested in the time-varying mechanical stress $\sigma(x,t)$ at location x from some reference point, and at time t. Following Korhonen's formulation, σ is positive for tensile stress and negative for compressive stress, and can be obtained by solving the PDE:

$$\frac{\partial \sigma}{\partial t} = \frac{B\Omega}{k_b T_m}\frac{\partial}{\partial x}\left\{D_a\left(\frac{\partial \sigma}{\partial x} - \frac{q^*\rho}{\Omega}j\right)\right\} \tag{1}$$

where D_a is the coefficient of atomic diffusion, B is the bulk modulus, Ω is the atomic volume, k_b is the Boltzmann's constant, T_m is the temperature in Kelvin, q^* is the absolute value of the effective charge of the conductor and ρ is the resistivity of the conductor. A void nucleates in the line once the stress exceeds a predefined threshold value $\sigma_{th} > 0$. As in most recent works on EM, we assume that diffusivity D_a is the same throughout a metal line. As a result, in our work voids will nucleate only at the end points of metal lines. This is a mild assumption [10] [9] because it is much more common in the field to find voids and line failures at the end points of metal branches. Another quantity of interest is the atomic flux in the line, J_a, defined as the number of atoms that cross a cross-section of the line per second, per unit area. The atomic flux can be written as [8] [4]:

$$J_a = \frac{D_a C\Omega}{k_b T_m}\left(\frac{\partial \sigma}{\partial x} - \frac{q^*\rho}{\Omega}j\right) \tag{2}$$

where C is the concentration of atoms. Note, J_a can be positive or negative, depending on the reference directions chosen and the actual direction of the electric current.

2.3 Diffusivity of metal lines

The atomic diffusion coefficient D_a is usually expressed using the Arrhenius law:

$$D_a = D_0 e^{-Q/(k_b T_m)} \tag{3}$$

where D_0 is a constant and Q is the activation energy for vacancy formation and diffusion. Due to randomness in the microstructure of a metal line, the failure time due to EM is a random variable. This randomness is primarily accounted for by the corresponding randomness in D_a, which depends on the angles between grain boundaries that are part of the micro-structure, and is assumed to be lognormally distributed [11] with mean D_{avg}. Strictly speaking, D_a also depends on the stress value at a given point. However, it has been reported that the numerical results with stress dependent D_a are "not too different" from constant D_a [8]. Hence, as in many previous works [7] [10], we will assume that D_a is stress-independent.

2.4 Method of Lines

The method of lines (MoL) is a special finite-difference technique for solving PDEs [14]. The basic idea of MoL is to discretize the PDE in *all but one* independent variable, so that we are left with a set of Ordinary Differential Equations (ODE) that approximate the PDE. We can then use well established methods to solve the ODE numerically.

Discretizing the PDE along any variable requires us to approximate the partial derivatives. For a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, the partial derivative with respect to the i^{th} variable x_i can be approximated using a *central difference formula* [14]:

$$\frac{\partial f}{\partial x_i}(x) \approx \frac{f(x + e_i\Delta x) - f(x - e_i\Delta x)}{2\Delta x} \tag{4}$$

$$\frac{\partial^2 f}{\partial x_i^2}(x) \approx \frac{f(x + e_i\Delta x) + f(x - e_i\Delta x) - 2f(x)}{(\Delta x)^2} \tag{5}$$

where Δx is a small positive scalar and e_i is the i^{th} unit vector (a vector that has 1 in position i and 0 elsewhere).

2.5 Limited Distributions

Let $\mathbf{Y}$ be a random variable (RV) with cumulative distribution function (cdf) $F_Y(t)$ and let l and u be two scalars with $l < u$ and at least one of them is finite. Then, the *limited RV* $\mathbf{Y'}$ between limits l and u has the following cdf [1]:

$$F_{Y'}(t) = \begin{cases} 0, & t < l \\ F_Y(t), & l \le t < u \\ 1, & t \ge u \end{cases} \tag{6}$$

In this paper, we refer to $\mathbf{Y}$ as the underlying RV and $\mathbf{Y'}$ as the limited RV.

3. INTERCONNECT TREE EM ANALYSIS

Modern power grids are made of Copper (Cu) and are fabricated using dual damascene process. In dual-damascene process, the metal line and via are formed simultaneously using copper. A barrier metal liner (usually Tantalum) must completely surround all Cu interconnects to prevent it from diffusing into the surrounding dielectric. Cross section of a typical metal via structure in Cu dual damascene process is as shown in Fig. 1. Each layer mostly consists of parallel stripes that are connected by vias

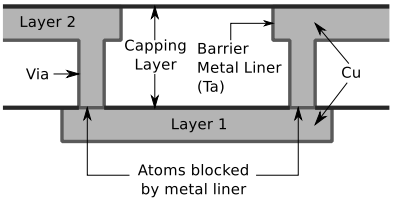

Figure 1: Cross sectional schematic of Cu dual damascene interconnects

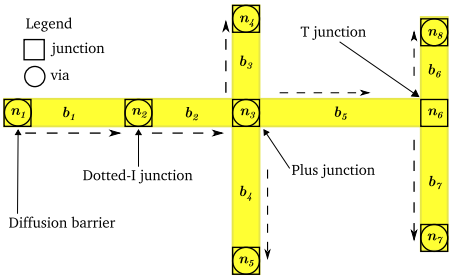

Figure 2: A typical interconnect tree structure.

to other layers. Note that due to the presence of barrier metal liner around vias between layers, Cu atoms from one layer cannot diffuse to another layer. On every layer, power and ground stripes are interspersed. As a result, the metal segments on every layer are mostly *trees*, i.e., they contain no loops or cycles. Thus, all previous work in this area assumes that the grid is made up of *interconnect trees*.

An interconnect tree is a continuously connected acyclic structure of straight metal lines within one layer of metalization such that atomic flux can flow freely within it. Fig. 2 shows a typical interconnect tree structure. Formally, an interconnect tree is a graph $\mathcal{T} = (\mathcal{N}, \mathcal{B})$ with no cycles, where $\mathcal{N}$ is a set of grid *junctions* and $\mathcal{B}$ is a set of resistive *branches*. A *branch* is defined to be a continuous straight metal line of uniform width. A *junction* is any point on the interconnect tree where a branch ends or where a via is located. Usually, but not always, current density around a junction is discontinuous. This discontinuity can be brought about by either a change in the width of connected branches, or by a change in the current itself due to the presence of a via. Thus, every via is a junction but the converse is not true. We define the *degree* of a junction to be equal to the number of branches connected to it. Note that a via does not contribute to the degree of a junction. In this paper, a junction with degree 1 will be referred to as a *diffusion barrier*, a junction with degree 2 will be referred to as a *dotted-I* junction, a junction with degree 3 will be referred to as a T junction and a junction with degree 4 will be referred to as a *Plus* junction. We treat corners in a tree as *dotted-I* junctions. Junctions with degrees higher than 4 are rarely found in practice. In general, it is worth noting that interconnect trees are *always* terminated by diffusion barriers and/or vias, hence the atoms cannot diffuse from one tree to another, and that different branches within a tree can have different widths.

3.1 Assigning reference directions

Before doing any analysis, we need to assign reference directions to all branches. This is necessary in order to consistently track the directions of branch currents and atomic flux.

An interconnect tree is equivalent to a graph, with grid junctions as vertices and branches as edges. With this analogy, there are many ways to assign reference direction to the branches. We choose the following way: starting from any diffusion barrier, we traverse the whole interconnect tree using a *breadth-first search* on the graph. This creates predecessor-successor relationships between the junctions. The reference direction for each branch is then assigned from predecessor to successor. The branch current (and atomic flux) is positive if it flows in the reference direction, otherwise it is negative. Likewise, the reference point for distance

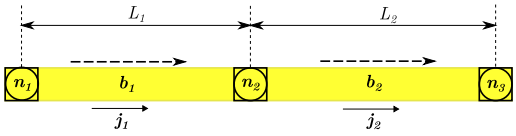

Figure 3: A simple 3-terminal tree $\mathcal{T}_d$.

is the predecessor junction, so that $x = 0$ is the predecessor and $x = L$ (line length) is the successor. In Fig. 2, if we start from the leftmost diffusion barrier (labeled as n_1), then the reference directions for each branch would be as shown by the dashed arrow lines.

3.2 Extending Korhonen's model to trees

In order to correctly estimate the level of EM degradation within a tree, we will extend Korhonen's model to account for the coupling between the branches. For better understanding, we illustrate our approach with a simple example. Consider a simple tree $\mathcal{T}_d = (\mathcal{N}, \mathcal{B})$, with $\mathcal{N} = \{n_1, n_2, n_3\}$ and $\mathcal{B} = \{b_1, b_2\}$, with reference directions as shown in Fig. 3. Branch b_k has dimensions $L_k \times w_k \times h_k$ (length × width × height), carries a current density j_k and has an atomic diffusivity of $D_{a,k}$, where $k = \{1, 2\}$. Note that $x_1 = L_1$ and $x_2 = 0$ denote the same point: the location of n_2. We are interested in stress as a function of position and time, i.e. $\sigma_1(x_1, t)$ and $\sigma_2(x_2, t)$) for branches b_1 and b_2, respectively. Once σ_1 and σ_2 are known, we can easily determine the EM degradation in the branches.

As we will see in Section 4, Korhonen's model (18) gives us the time rate of change of stress for a point *within* a branch, for $k = \{1, 2\}$, as follows:

$$\frac{\partial \sigma_k}{\partial t} = \frac{B\Omega}{k_b T_m} \frac{\partial}{\partial x_k} \left\{ D_{a,k} \left(\frac{\partial \sigma_k}{\partial x_k} - \frac{q^* \rho}{\Omega} j_k \right) \right\}, \; x_k \in (0, L_k) \quad (7)$$

However, in order to solve the PDE for the whole tree, we need to also state the boundary conditions at all end-points of branches, i.e. at junctions. The boundary conditions describe the behaviour of stress at the junctions. For the example in Fig. 3, we will discuss the two cases of a diffusion barrier and a dotted-I junction.

3.2.1 Diffusion Barrier

Junctions n_1 and n_3 are diffusion barriers, where the atomic flux is blocked. Considering the nucleation phase first, J_a is zero at the barrier so that from (2):

$$\begin{aligned} J_{a,1}(0,t) = 0 \quad &\Longrightarrow \quad \frac{\partial \sigma_1(0,t)}{\partial x_1} = \frac{q^* \rho}{\Omega} j_1 \\ J_{a,2}(L_2,t) = 0 \quad &\Longrightarrow \quad \frac{\partial \sigma_2(L_2,t)}{\partial x_2} = \frac{q^* \rho}{\Omega} j_2 \end{aligned} \quad (8)$$

We next move to the void growth phase. For a void to nucleate at n_1 (n_3), the electronic current must move away from n_1 (n_3), so that $j_1 < 0$ ($j_2 > 0$). Exactly what happens around a void is somewhat complicated and cannot really be captured in a 1D model. The recent work in [15] provides an extension of the Korhonen 1D model to describe behaviour of stress around a void. From this, stress falls to zero at the void surface but remains at its original value a very short distance $\delta \approx 1$nm from the void surface. We refer to δ as the *thickness of the void interface*. From [15], the stress gradients at junctions n_1 and n_3 throughout the void growth phase are:

$$\frac{\partial \sigma_1(0,t)}{\partial x_1} = \frac{\sigma_1(0,t)}{\delta} \quad \text{and} \quad \frac{\partial \sigma_2(L_2,t)}{\partial x_2} = -\frac{\sigma_2(L_2,t)}{\delta} \quad (9)$$

where $\sigma_1(0,t) = \sigma_2(L_2,t) = \sigma_{th}$ at the time of void nucleation.

3.2.2 Dotted-I Junction

The atomic flux interaction at dotted-I junction n_2 is the key to describing the coupling of stresses in branches b_1 and b_2. Considering the nucleation phase first, the metal is continuous across n_2, which is the same physical point of both b_1 and b_2, so that:

$$\sigma_1(L_1, t) = \sigma_2(0, t) \quad (10)$$

and atomic flux can flow freely between b_1 and b_2 [6]. Because the material flow across an infinitesimal boundary at n_2 has to be continuous, we have:

$$w_1 h_1 J_{a,1}(L_1, t) = w_2 h_2 J_{a,2}(0, t) \tag{11}$$

Next considering the void growth phase, once a void nucleates at n_2, it is shared by both branches b_1 and b_2. For our 1D model, we make the reasonable assumption that the void covers the entire cross-sectional area of the junction. As a result, there would be no flow of atomic flux between b_1 and b_2. Hence, during the *void growth* phase, we effectively treat n_2 as a diffusion barrier for both branches b_1 and b_2, so that:

$$\frac{\partial \sigma_1(L_1, t)}{\partial x_1} = -\frac{\sigma_1(L_1, t)}{\delta} \quad \text{and} \quad \frac{\partial \sigma_2(0, t)}{\partial x_2} = \frac{\sigma_2(0, t)}{\delta} \tag{12}$$

Using (7) and the boundary conditions obtained from (8), (9) and (10)-(12), we then formulate an *initial value problem* (IVP) to solve for stress. As we will see later, the IVP combined with the initial condition $\sigma_1(x_1, 0) = \sigma_2(x_2, 0) = \sigma_{res}$, where σ_{res} is the so-called residual stress [7] before the application of electric current, will completely determine σ_1 and σ_2. We will next generalize the above schemes for capturing flux interactions at junctions, into a set of *laws* that forms the basis for our approach.

3.3 Boundary Laws for junctions

Boundary laws govern the interaction of atomic flux at junctions. Consider a junction n_p, and let $\mathcal{B}_p$ be the set of branches connected to n_p. Let $t_{f,p}$ be the time of void nucleation for this junction. Then, the boundary laws (motivated mainly by the law of conservation of mass) can be stated as:

Law 1. For $t < t_{f,p}$, the number of metal atoms flowing into n_p per unit time is the same as the number of metal atoms flowing out from it:

$$\sum_{b_k \in \mathcal{B}_{p,in}} w_k h_k J_{a,k} = \sum_{b_k \in \mathcal{B}_{p,out}} w_k h_k J_{a,k} \tag{13}$$

where w_k (h_k) is the width (height) of the branch, $\mathcal{B}_{p,in}$ is the set of branches for which the reference direction is going into n_p, and $\mathcal{B}_{p,out}$ is the set of branches for which the reference direction is going out from n_p.

Law 2. For $t \geq t_{f,p}$, there is no flow of atomic flux between the connected branches $\mathcal{B}_p$. The stress gradient at the junction, generalizing from (9) and (12), is:

$$\frac{\partial \sigma_{k,p}}{\partial x_k} = \pm \frac{\sigma_{k,p}}{\delta} \tag{14}$$

where $\sigma_{k,p}$ is the value of stress at end-point n_p of branch b_k. The sign is positive for $b_k \in \mathcal{B}_{p,out}$ and negative for $b_k \in \mathcal{B}_{p,in}$.

Law 3. Until a void nucleates at n_p, the stress values in any two branches where they meet at n_p are equal.

3.4 Void growth and resistance change

Once the stress at any point in the tree reaches σ_{th}, a void nucleates at that point. As noted before, in our EM model, void nucleation can occur only at junctions and not within the branches. We assume that once a void nucleates at a junction, it is shared by all the branches connected to that junction. Tracking void growth is useful in order to determine the branch resistances and the corresponding current densities. Recent work [15] shows that the initial void growth rate is very high. Hence, as a conservative approximation, we assume that once a void nucleates at any junction n_p, the void lengths for all branches b_k connected to n_p reach their steady state values in a very short period of time. Thus, when a void nucleates we assume that line resistance rises immediately to its steady state value, for all connected branches. The steady state value for void length is:

$$l_{k,v} = L_k \left(\frac{\sigma_T}{B} + \frac{q^* \rho j_k L_k}{2B\Omega} \right) \tag{15}$$

and correspondingly, the branch resistance will be:

$$R_k = \rho_b l_{k,v} / A_b + \rho_m (L_k - l_{k,v}) / A_m \tag{16}$$

where σ_T is the thermal stress [15], and $\rho_m(\rho_b)$ and $A_m(A_b)$ are the resistivity and cross-sectional area of the metal (liner), respectively. Note that for any branch b_k, $l_{k,v}$ and j_k are interdependent on each other. As such, we iteratively find j_k and $l_{k,v}$ using modified Richardson iteration.

4. SOLVING THE EXTENDED MODEL

In this section, we will describe our approach for solving the extended Korhonen's model for trees. First, for points within a branch, we will use the method of lines (MoL) to convert the PDEs into a set of ODEs by discretizing along the spatial domain. Then, using the laws proposed in Section 3.3, we will derive the boundary conditions at the junctions. Finally, we merge the two and state the final initial value problem (IVP) that describes the stress evolution for a given tree.

4.1 Scaling Korhonen's model

Korhonen's model (1) is often scaled by introducing dimensionless variants of stress, distance (length) and time [4]. This leads to stable PDEs that are easier to solve numerically. We define the following scaling factors for any branch $b_k \in \mathcal{B}$:

$$\tau \triangleq \frac{B\Omega}{k_b T_m} \frac{D_{avg} t}{L_c^2}, \quad \eta_k \triangleq \frac{\Omega \sigma_k}{k_b T_m}, \quad \xi_k \triangleq \frac{x_k}{L_k} \tag{17}$$

where D_{avg} is the average atomic diffusivity of the metal (see Section 2.3), L_c is some chosen characteristic length to ensure proper scaling and $0 \leq x_k \leq L_k$. The new variables τ, η and ξ are referred to as reduced time, stress and distance, respectively. Using (17) in (1) and applying the chain-rule, we get:

$$\frac{\partial \eta_k}{\partial \tau} = \theta_k \frac{\partial}{\partial \xi_k} \left(\frac{\partial \eta_k}{\partial \xi_k} - \alpha_k \right) \tag{18}$$

where $\theta_k = (L_c^2 D_{a,k})/(L_k^2 D_{avg})$, $\alpha_k = (q^* \rho j_k L_k)/(k_b T_m)$, j_k is the current density and $D_{a,k}$ is the diffusivity for b_k. Since, for any given branch, α_k is not a function of distance ξ_k, then $\partial \alpha_k / \partial \xi_k = 0$ and we get:

$$\frac{\partial \eta_k}{\partial \tau} = \theta_k \frac{\partial^2 \eta_k}{\partial \xi_k^2} \tag{19}$$

For any branch b_k, (19) constitutes the scaled PDE system to be solved. Also, the atomic flux in b_k can be restated in terms of reduced variables:

$$J_{a,k} = \frac{D_{a,k} C}{L_k} \left(\frac{\partial \eta_k}{\partial \xi_k} - \alpha_k \right) \tag{20}$$

4.2 Discretization for a tree branch

We uniformly discretize branch b_k into N segments, where N is the same for all branches (because we have scaled all branch lengths to 1 as in (17)). The reduced stress at each of the $N+1$ discrete spatial points $\{0, 1, \ldots N\}$ in the branch is denoted by $\eta_{k,i}$ and the time rate of change of $\eta_{k,i}$ is (from (19)):

$$\frac{\partial \eta_{k,i}}{\partial \tau} = \theta_k \frac{\partial^2 \eta_{k,i}}{\partial \xi_k^2} \quad \text{for } i = 0, 1, \ldots N \tag{21}$$

Further, we approximate the partial derivatives with respect to ξ using the central difference approximation, so that (21) leads to:

$$\frac{\mathrm{d}\eta_{k,i}}{\mathrm{d}\tau} = \theta_k \left(\frac{\eta_{k,i+1} + \eta_{k,i-1} - 2\eta_{k,i}}{(\Delta\xi)^2} \right) \tag{22}$$

where $\Delta\xi = \Delta\xi_k = 1/N$, $\forall k$. The corresponding atomic flux $J_{a,k,i}$ at the i^{th} point is:

$$J_{a,k,i} = \frac{D_{a,k} C}{L_k} \left(\frac{\eta_{k,i+1} - \eta_{k,i-1}}{2\Delta\xi} - \alpha_k \right) \tag{23}$$

Note that for each branch, the ODEs at junctions ($i = \{0, N\}$) require the values for $\eta_{k,-1}$ and $\eta_{k,N+1}$, which are *not* part of the ξ_k domain. The values at these *ghost points* are obtained by solving for the respective boundary condition(s), as we next explain.

4.3 Boundary Conditions at diffusion barrier

Consider a diffusion barrier n_p connected to branch b_k. We have two cases, one where n_p is at the predecessor junction ($\xi_k = 0$, start of the branch) and one where it is at the successor junction ($\xi_k = 1$, branch end). We first obtain the boundary conditions for n_p at $\xi_k = 0$. Let τ_f be the time of void nucleation at this barrier. Then, the corresponding boundary condition is (using (13) and (14)):

$$\frac{\partial \eta_{k,0}}{\partial \xi_k} = \begin{cases} \alpha_k & \tau < \tau_f, \\ \eta_{k,0}(L_k/\delta) & \tau \geq \tau_f \end{cases} \tag{24}$$

where $\eta_{k,0}$ corresponds to $\sigma_{k,p}$ in (14), with $\eta_{k,0} = \eta_{th}$ at $\tau = \tau_f$. Using the central difference approximation, we get:

$$\eta_{k,-1} = \begin{cases} \eta_{k,1} - 2\Delta\xi\alpha_k & \tau < \tau_f, \\ \eta_{k,1} - 2\Delta\xi\eta_{k,0}(L_k/\delta) & \tau \geq \tau_f \end{cases} \tag{25}$$

Similarly, for a diffusion barrier at $\xi_k = 1$, we get:

$$\eta_{k,N+1} = \begin{cases} \eta_{k,N-1} + 2\Delta\xi\alpha_k & \tau < \tau_f, \\ \eta_{k,N-1} - 2\Delta\xi\eta_{k,N}(L_k/\delta) & \tau \geq \tau_f \end{cases} \tag{26}$$

4.4 Boundary Conditions at dotted-I junction

Consider a dotted-I junction n_p. Without loss of generality, we will assume that n_p is at the end of branch 1 and at the beginning of branch 2. To solve the ODE at n_p, we need the value of at least one of the ghost points ($\eta_{1,N+1}$ or $\eta_{2,-1}$). Let τ_f be the time of void nucleation at this junction. Then, using (13), we get ($h_1 = h_2$ within a metal layer):

$$w_1 J_{a,1,N} - w_2 J_{a,2,0} = 0 \quad \text{for } \tau < \tau_f \tag{27}$$

Also, from law 3, $\eta_{1,N} = \eta_{2,0}$ when $\tau < \tau_f$. Hence, the time rate of change of stress should also be the same, so that using (19):

$$\frac{\partial \eta_{1,N}}{\partial \tau} = \frac{\partial \eta_{2,0}}{\partial \tau} \implies \frac{\partial^2 \eta_{1,N}}{\partial \xi_1^2} = \frac{\theta_2}{\theta_1}\frac{\partial^2 \eta_{2,0}}{\partial \xi_2^2} \quad \text{for } \tau < \tau_f \tag{28}$$

Substituting the value of J from (23) in (27) and applying the central difference formula in (28), we can obtain the value of ghost points. Due to lack of space, we omit the complete derivation and only present the final value of $\eta_{1,N+1}$:

$$\eta_{1,N+1} = \eta_{1,N-1} + (r_{12}q_1 + w_{21}q_2)/(r_{12} + w_{21}) \tag{29}$$

where $r_{12} = L_1/L_2$, $w_{21} = w_2/w_1$, $p_{21} = D_{a,2}/D_{a,1}$, $q_1 = 2\Delta\xi\,(\alpha_1 - w_{21}r_{12}p_{21}\alpha_2)$ and $q_2 = 2(r_{12}^2 p_{21}\eta_{2,1} - \eta_{1,N-1} + (1 - r_{12}^2 p_{21})\eta_{1,N})$. Using law 2, n_p is treated as a diffusion barrier during the void growth phase. Thus, for $\tau \geq \tau_f$:

$$\eta_{1,N+1} = \eta_{1,N-1} - 2\Delta\xi\eta_{1,N}(L_1/\delta)$$
$$\eta_{2,-1} = \eta_{2,1} - 2\Delta\xi\eta_{2,0}(L_2/\delta)$$

Due to space constraints, we are also unable to show the corresponding boundary conditions for T and *plus* junctions, but they can be obtained by following the same procedure as done for the dotted-I junction.

4.5 Formulating the IVP for a tree

For any given tree $\mathcal{T} \in \{\mathcal{N}, \mathcal{B}\}$, the discretized equations (21) for branches $b_k \in \mathcal{B}$, combined with initial and boundary conditions give the corresponding IVP formulation. Again, for clarity, we will revisit the example presented in Section 3.2. For simplicity, we will show the IVP formulation only for the scenario when no voids are present in the tree. For $\mathcal{T}_d$ (shown in Fig. 3) we can describe the stress evolution in terms of reduced variables as:

$$\text{PDE} \quad \frac{\partial \eta_{k,i}}{\partial \tau} = \theta_k \frac{\partial^2 \eta_{k,i}}{\partial \xi_k^2} \quad \text{for } i = \{0, 1, \ldots N\},\ k = \{1, 2\}$$
$$\text{B.C.} \quad J_{a,1,0} = 0,\ J_{a,2,N} = 0,\ w_1 J_{a,1,N} - w_2 J_{a,2,0} = 0, \text{ and } \eta_{1,N} = \eta_{2,0}$$
$$\text{I.C.} \quad \eta_{k,i} = \eta_{res} \text{ at } \tau = 0 \quad \forall k, i$$

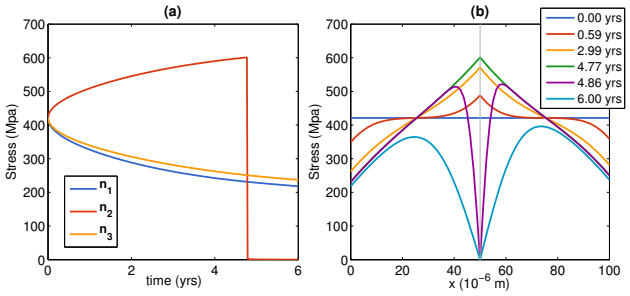

Figure 4: For $\mathcal{T}_d$, (a) stress evolution at junctions with time and (b) stress profile with time. $L_1 = L_2 = 50\ \mu m$, $j_1 = -j_2 = 6e9\ A/m^2$ and $\sigma_{th} = 600\ MPa$.

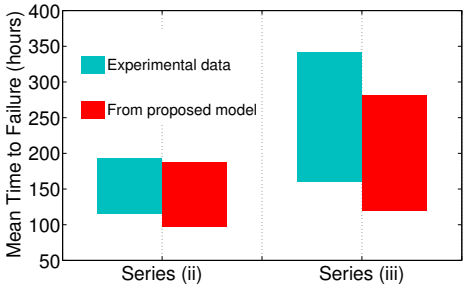

Figure 5: 95% confidence bounds on MTFs as estimated by the Extended Korhonen's model vs the experimental results as reported in Fig. 3 of Gan et. al. [5].

where $\eta_{res} = (\Omega\sigma_{res})/(k_b T_m)$ is the reduced initial stress, called the residual stress [7]. Then, using (22), (25), (26) and (29) in the above equations, we obtain the following ODEs:

$$\begin{aligned}
\mathrm{d}\eta_{1,0}/\,\mathrm{d}\tau &= -K_1\,(2\eta_{1,0} - 2\eta_{1,1}) + K_1 q_{n_1} \\
\mathrm{d}\eta_{1,i}/\,\mathrm{d}\tau &= K_1\,(\eta_{1,i-1} - 2\eta_{1,i} + \eta_{1,i+1}) \quad i \in \{1, N-1\} \\
\mathrm{d}\eta_{1,N}/\,\mathrm{d}\tau &= \varrho_{12}K_1\big(2\eta_{1,N-1} - 2(1 + p_{21}r_{12}w_{21})\eta_{1,N} \\
&\quad + 2p_{21}r_{12}w_{21}\eta_{2,1}\big) + \varrho_{12}K_1 q_{n_2} (= \mathrm{d}\eta_{2,0}/\,\mathrm{d}\tau) \\
\mathrm{d}\eta_{2,i}/\,\mathrm{d}\tau &= K_2(\eta_{2,i-1} - 2\eta_{2,i} + \eta_{2,i+1}) \quad i \in \{1, N-1\} \\
\mathrm{d}\eta_{2,N}/\,\mathrm{d}\tau &= K_2\left(2\eta_{2,N-1} - 2\eta_{2,N}\right) + K_2 q_{n_3}
\end{aligned} \tag{30}$$

where

$$q_{n_1} = -2\Delta\xi\alpha_1,\ q_{n_2} = 2\Delta\xi(\alpha_1 - w_{21}r_{12}p_{21}\alpha_2),\ q_{n_3} = 2\Delta\xi\alpha_2$$
$$K_1 = \theta_1/(\Delta\xi)^2,\ K_2 = \theta_2/(\Delta\xi)^2,\ \varrho_{12} = r_{12}/(r_{12} + w_{21})$$

Formulating the IVP for any given tree can be done by following the same procedure. We use *adaptive step Runge-Kutta* methods to solve (30), and the solution for $\mathcal{T}_d$ is as shown in Fig. 4.

4.6 Verifying the model

Fig. 5 shows a comparison of the MTF as obtained using the extended Korhonen's model and the experimental results reported in Fig. 3 of Gan et. al. [5]. The model parameters used are same as given in [5]. The comparison is made in terms of 95% confidence interval on the MTF, and as can be seen from Fig. 5, there is a good agreement between the results of our model and the experimental data. Clearly, by accounting for material flow between connected branches, the extended Korhonen's model can be used for determining the EM degradation in interconnect trees.

Now that we have presented and verified the *Extended Korhonen's model*, we will describe how it can be used for EM assessment of the entire power grid.

5. POWER GRID EM ANALYSIS

Because EM is a long-term failure mechanism, short-term transients that may be typically experienced in chip workloads do not play a significant role in EM degradation. Hence, and consistent with standard practice in the field, we use an effective-current model [16], so that the grid currents are assumed to be constant at some average (effective) value, at least during the void nucleation phase. Once a void nucleates, branch resistances change fairly quickly and the currents change, also fairly quickly, to new effective values. Thus, between any two successive void nucleations, the grid has fixed currents, voltages, and conductances and so can be modeled using a DC model. To denote the fact that conductances (and the corresponding voltages) change from one nucleation phase to the next, as in [3], we express the grid model as:

$$G(t)v(t) = i_s \tag{31}$$

where $G(t)$ is the time-varying (but piecewise-constant) conductance matrix, $v(t)$ is the corresponding time-varying (but piecewise constant) vector of node voltage drops and i_s is the vector of effective values of the current sources tied to the grid.

5.1 The Main Approach

As explained earlier, we use the *mesh model* to find the Mean Time to failure (MTF), in which the grid is deemed to have failed when enough voids have nucleated that the voltage drop specification exceeds at some grid node. As a byproduct, however, this process also produces the time when the first void nucleates, which helps us generate the MTF as per the *series model*, in which a grid is deemed to have failed when the first void nucleates. We report the series model MTF for comparison purposes.

We assume that the grid is undamaged (no voids) at $t = 0$. A voltage-drop threshold value for every grid node (or a subset of grid nodes) is given, which is captured in the vector v_{th}. Initially, all node voltage drops are less than v_{th}, i.e. $v(0) < v_{th}$. A power grid is a collection of interconnect trees. As such, to estimate the EM degradation of the grid, we formulate an IVP for *every tree* as shown in Section 4.5 and numerically integrate it to obtain the stress as a function of position and time. Every time a void nucleates at a junction, we calculate the new branch resistances and current density values. The corresponding boundary conditions are updated and we re-formulate the IVP for all trees using the updated boundary conditions. The time of first void nucleation gives the time to failure (TTF) of the grid as per the series model. Due to increase in branch resistances, the voltage drops in the grid continue to increase as we move forward in time. The earliest time when the voltage drop exceeds v_{th} is the TTF of the grid as per the mesh model.

To account for the random nature of EM degradation, we perform *Monte Carlo random sampling* to estimate the MTF. In each Monte Carlo iteration, we assign new randomly-generated diffusivity values to all the branches in the grid. This effectively produces a new *instance* of the whole power grid, which we refer to as a *sample grid*. We then use the above extended Korhonen model and the IVP formulation to generate a TTF value based on the series model and another that is based on the mesh model. With enough samples, we form two averages as our estimates of the series MTF and the mesh MTF.

Let $\mathbf{T}$ be the RV that represents the statistics of mesh grid TTF for this approach, then expected value of $\mathbf{T}$, denoted by $E[\mathbf{T}]$, is the mesh MTF of the grid. Using goodness of fit methods, we found that normal distribution is a good fit for $\mathbf{T}$ (see Fig. 6). Therefore, we can use standard statistical sampling (Monte Carlo) [12] to find the value of $E[\mathbf{T}]$ to within user-specified error tolerance. The number of samples required for Monte Carlo to terminate is determined such that we have $(1-\lambda)\times 100\%$ confidence (e.g. $\lambda = 0.05$ for 95% confidence) that the relative error in MTF estimation is less than a user provided relative error threshold ϵ (e.g. $\epsilon = 0.1$ for 10% relative error threshold).

Though this is the most accurate approach, numerically solving *all* the trees in the power grid using the extended Korhonen's model is computationally expensive. Hence, we use the main approach only on smaller grids. The results from this approach serve as a benchmark of comparison for more optimized approaches.

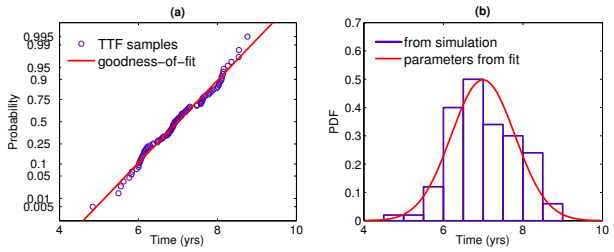

Figure 6: (a) Goodness-of-fit plot for normal distribution and (b) pdf using 100 mesh TTF samples from ibmpg2 *main* approach.

5.2 Improved performance

We will now present a method that drastically reduces the runtime with almost no impact on accuracy of results. We will refer this as the *Filtering* approach. For each *sample grid*, solving all the trees up to the time of grid failure yields a specific sequence of void nucleation times in certain trees that are of interest. In particular, all trees that nucleate their first void *before* the time of grid failure are of interest to us. All trees that nucleate their first void *after* the grid failure are inconsequential to us, and we would do well to not solve them in the first place. Unfortunately, we don't know up-front which set of trees should be solved, and which can be discarded. However, we can devise an approximate but conservative filtering scheme that indicates which subset of trees will most likely nucleate before all the rest, which we will call the *active set*, and we then solve only these trees. To do this, we use the easy-to-compute but approximate Riege and Thompson model [6] to *estimate* the time to first void nucleation in a tree. We have verified empirically that this is always less than or equal the true time to first void nucleation, thus giving us a conservative estimate of the subset of trees leading to grid failure.

For a given sample grid, we restrict our attention to trees whose estimated first void nucleation times are smaller than some threshold $t = t_m$. We refer to t_m as the *active set cutoff threshold*. t_m is a part of the Monte Carlo process. We start with a sufficiently high value for t_m, that is reduced as more samples are obtained. Trees that are likely to nucleate before t_m, based on the Riege and Thompson estimate, are declared to be part of the active set and will be numerically solved. Because the estimate is conservative, then the active set may include some trees that ideally would not need to be solved. If the *sample grid* fails before t_m, we obtain a sample TTF. On the other hand, if the sample grid hasn't failed up to $t = t_m$, we set the TTF sample equal to t_m, and such a sample is called a *limited sample*. Thus, in our Monte Carlo approach, we effectively sample from RV $\mathbf{T}'$ that has a *limited normal distribution* ($l = -\infty$ and $u = t_m$) with underlying normal RV $\mathbf{T}$. However, we can estimate the mean of the underlying RV $E[\mathbf{T}]$ using the samples obtained from the limited RV $\mathbf{T}'$, as shown next.

For both RV's $\mathbf{T}$ and $\mathbf{T}'$, using the *law of total expectation* [17], we can write:

$$E[\mathbf{T}] = E[\mathbf{T}|\mathbf{T} \le t_m]F_T(t_m) + E[\mathbf{T}|\mathbf{T} > t_m](1 - F_T(t_m)) \tag{32}$$
$$E[\mathbf{T}'] = E[\mathbf{T}'|\mathbf{T}' \le t_m]F_{T'}(t_m) + E[\mathbf{T}'|\mathbf{T}' > t_m](1 - F_{T'}(t_m))$$

where $F_T(t)$ is the cdf of the normal RV $\mathbf{T}$ and $F_{T'}(t)$ is the cdf of limited normal RV $\mathbf{T}'$. From the definition of limited RV, we have $E[\mathbf{T}'|\mathbf{T}' \le t_m] = E[\mathbf{T}|\mathbf{T} \le t_m]$, $E[\mathbf{T}'|\mathbf{T}' > t_m] = t_m$ and $F_{T'}(t_m) = F_T(t_m)$. Hence, we can write:

$$E[\mathbf{T}'] = E[\mathbf{T}|\mathbf{T} \le t_m]F_T(t_m) + t_m(1 - F_T(t_m)) \tag{33}$$

Subtracting (33) from (32), we get:

$$\begin{aligned} E[\mathbf{T}] &= E[\mathbf{T}'] + (E[\mathbf{T}|\mathbf{T} > t_m] - t_m)(1 - F_T(t_m)) \\ &= E[\mathbf{T}'] + E[\mathbf{T} - t_m|\mathbf{T} > t_m](1 - F_T(t_m)) \end{aligned} \tag{34}$$

The term $E[\mathbf{T} - t_m|\mathbf{T} > t_m]$ is the Mean Residual Life (MRL) of the power grid at $t = t_m$. Define $\mu \triangleq E[\mathbf{T}]$, $\mu' \triangleq E[\mathbf{T}']$ and

$p_f \triangleq F_T(t_m)$. Since we know $\mathbf{T}$ has a normal distribution, the MRL of the power grid at $t = t_m$ can be expressed in terms of μ and p_f. From (34), after some algebraic manipulation, we obtain:

$$\mu = \frac{\mu' + (\kappa - 1)t_m}{\kappa} \tag{35}$$

where $\kappa = p_f + \phi\left(\Phi^{-1}(p_f)\right)/\Phi^{-1}(p_f)$, $\Phi(t)$ and $\phi(t)$ are respectively the cdf and probability distribution function (pdf) of a standard normal distribution $\mathcal{N}(0,1)$. Φ^{-1} denotes the inverse cdf of of $\mathcal{N}(0,1)$ which can be computed on most operating systems using `erfinv()` function. As we will show next, we can estimate μ' and p_f from the statistical sampling process. Thus, we can evaluate the RHS of (35) to estimate μ.

Similar to the *main* approach, we stop the Monte Carlo process when we are $(1-\lambda)\times 100\%$ confident that the relative error in estimated MTF is less than some user provided threshold ϵ. In other words, we stop if:

$$\delta\mu_\lambda/\mu \leq \epsilon \iff \delta\mu_\lambda/\hat{\mu} \leq \epsilon/(1+\epsilon) \tag{36}$$

where $\delta\mu_\lambda$ is $(1-\lambda)\times 100\%$ confidence bound on the estimation error $\delta\mu$, as explained below.

Let $\{T'_1, T'_2, \ldots T'_s\}$ be s samples obtained from RV $\mathbf{T}'$ using a Monte Carlo process. Then, $\hat{\mu}' = (1/s)\sum_{k=1}^{s} T_k$ is the estimated value of μ' and $\hat{p}_f = 1 - |\{T'_k : T'_k > t_m\}|/s$ is the estimated value of p_f. Using $\hat{\mu}'$ and $\hat{p}_f$ in (35) we get $\hat{\mu}$, the estimated value of μ. Note that μ', p_f and μ are the *true* values, so that $\lim_{s\to\infty}\hat{\mu}' = \mu'$, $\lim_{s\to\infty}\hat{p}_f = p_f$ and $\lim_{s\to\infty}\hat{\mu} = \mu$. Then, the error in estimation is $\delta\mu = |\hat{\mu} - \mu|$, $\delta\mu' = |\hat{\mu}' - \mu'|$ and $\delta p_f = |\hat{p}_f - p_f|$. A $(1-\lambda)\times 100\%$ confidence bound on error $\delta\mu$ means that the interval $[\hat{\mu} - \delta\mu, \hat{\mu} + \delta\mu]$ will contain the true value $(1-\lambda)\times 100\%$ of the time.

Using the obtained TTF samples, we can find the confidence bounds on $\delta\mu'$ and δp_f (for a chosen confidence level λ), which in turn gives us the confidence bound on $\delta\mu$. For lack of space, we skip the details and present the final expression:

$$\delta\mu_\lambda = \sqrt{\frac{(\delta\mu'_\lambda)^2}{\hat{\kappa}^2} + \frac{z^2_{\lambda/2}(t_m - \hat{\mu}')^2\hat{p}_f(1-\hat{p}_f)}{\hat{\kappa}^4 s}\left[1 + \left(1 + \frac{1}{y^2}\right)^2\right]} \tag{37}$$

where $\delta\mu'_\lambda$ is the $(1-\lambda)\times 100\%$ confidence bound on μ', $z_{\lambda/2}$ is the $(1-\lambda/2)$-percentile of $\mathcal{N}(0,1)$, $\hat{\kappa} = \hat{p}_f + \phi(y)/y$ and $y = \Phi^{-1}(\hat{p}_f)$. Note $\delta\mu'_\lambda$ is obtained from simulation, using the technique given in [1]. We obtain at least 30 TTF samples before starting to check the stopping criteria (36).

5.2.1 TTF predictor approach

We next describe a predictor-based approach to further speed up the MTF computation. This approach is applied on top of the filtering approach explained earlier.

All test data that we've seen shows a profile of tensile stress evolution over time that follows a particular trend: it is a monotonically increasing function of time with a gently decreasing slope (see for example Fig. 4). Once a tree in the active set has been solved for a few time-points, it should be possible to extrapolate the rest of the trend, with some suitably nonlinear fitting function. The fitting function can thus be used as a *TTF predictor*, to find a good estimate of the nucleation times for all junctions within the tree. Parameters of the function can be found using least-squares fitting, based on the points already solved. While various exponential or log functions may be suitable, we have found empirically that the following power function template provides a very good fit:

$$f(t) = at^{b + c\ln t} \tag{38}$$

where a, b and c are parameters to be determined using regression analysis and least-squares fitting and $f(t)$ is the stress value at time t. Note that $\ln(f(t))$ is a simple quadratic in $\ln t$, with $\ln a$, b and c as the three coefficients. Once we estimate the time of void nucleations using the TTF predictor, we can predict the time (and sequence) of void nucleations. For each void nucleation, we update the corresponding branch resistances and voltage drops until the grid fails.

Table 1: Comparison of power grid (mesh) MTF using the *main* approach and *Filtering* approach

Grid Name	Main		Filtering		Error (%)	Speed-up
	$\mu_m^{\mathtt{all}}$ (yrs)	runtime (hrs)	$\mu_m^{\mathtt{act}}$ (yrs)	runtime (hrs)		
ibmpg1	6.14	0.15	6.10	0.08	0.62	1.92x
ibmpg2	6.65	2.10	6.63	0.87	0.28	2.41x
ibmpg5	5.98	20.51	5.94	0.53	**0.72**	**38.83x**

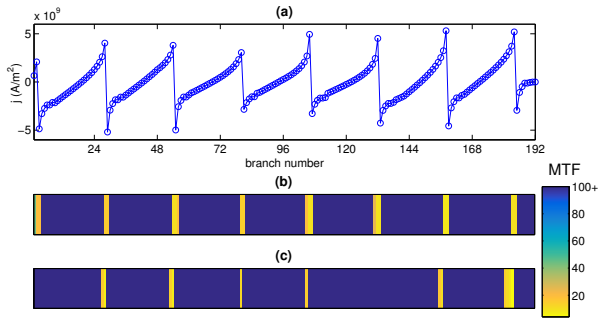

Figure 7: (a) High current density profile and possible failure locations as predicted using (b) Extended Korhonen's model and (c) Black's model

6. EXPERIMENTAL RESULTS

All approaches have been implemented in C++ and tested on a number of IBM power grid benchmarks [13], using a quad-core 3.4GHz Linux machine with 32GB of RAM. The *voltage drop threshold* v_{th} is defined to be 5% of v_{dd} for *all* grid nodes, except ibmpg6, for which v_{th} is defined to be 10% of v_{dd}. The interconnect material is assumed to be Copper, so that the following parameters are used in our EM model: $\sigma_{res} = 421\times 10^6$Pa [7], $B = 1.35\times 10^{11}$Pa, $\Omega = 1.66\times 10^{-29}\text{m}^3$, $k_b = 1.38\times 10^{-23}$J/K, $q^* = 8.0109\times 10^{-19}$C, $\sigma_{th} = 600\times 10^6$Pa [7] and $\delta = 10^{-9}$m. A nominal temperature of 373K is used for all simulations. Each branch is discretized into $N = 16$ segments. We use a relative tolerance of 10^{-3} and an absolute tolerance of 10^{-6} for the ODE solver. For all grids, we used $\lambda = 0.05$ (95% confidence bounds) and $\epsilon = 0.1$ (maximum relative error threshold of 10%). In our implementation, we use a *shared memory model* to parallelize the computation and take advantage of the quad core machine.

Table 1 compares the accuracy and runtime of the *main* approach vs. the *filtering* approach. Since the *main* approach slows down considerably as the grid size increases, we were able to use it on only the three smallest benchmarks. $\mu_m^{\mathtt{all}}$ and $\mu_m^{\mathtt{act}}$ denote the estimated MTF using the the main approach and the filtering approach, respectively. From table 1, it is clear that as the grid size increases, the filtering approach leads to significant speed-ups with negligible loss in accuracy. For the largest grid we could test (ibmpg5 with 2002 trees), the filtering approach obtained a speed-up of $\sim$ **39x** over the *main* approach with error being only 0.72%. This shows the value of the filtering approach.

Table 2 lists the MTFs estimated per the series and mesh models using three approaches, based on: *i)* Black's model as implemented in [3] (columns $\mu_s^{\mathtt{blk}}$ and $\mu_m^{\mathtt{blk}}$), *ii)* our extended Korhonen's model with the active set filter ($\mu_s^{\mathtt{act}}$ and $\mu_m^{\mathtt{act}}$) and *iii)* our extended Korhonen's model with the active set filter and the TTF predictor ($\mu_s^{\mathtt{pre}}$ and $\mu_m^{\mathtt{pre}}$). In order to give a fair comparison, we calibrate Black's model with data obtained from Korhonen's model. From the table, we note that $\mu_s^{\mathtt{act}} > \mu_s^{\mathtt{blk}}$ and $\mu_m^{\mathtt{act}} > \mu_m^{\mathtt{blk}}$ for all grids except ibmpg1. Overall, the mesh MTF estimated using our approach is 3x longer than that found using Black's model. Finally, in the table we compare the MTF results from the filter-only approach to those based on using the filter *and* TTF predictor. Overall, for all the reported grids, the predictor based variation of the filtering approach achieves an average speed-up of nearly 2x over the filter-only approach, with average error being 7.92%.

Table 2: Comparison of power grid MTF as estimated using Black's model and Extended Korhonen's model

Power Grid				Black's Model			Extended Korhonen's Model								
							Filtering			Filtering + Predictor					
Grid Name	# nodes	#bra-nches	#trees	$\mu_s^{\mathtt{blk}}$ (yrs)	$\mu_m^{\mathtt{blk}}$ (yrs)	$t_{\mathtt{blk}}$ (hrs)[a]	$\mu_s^{\mathtt{act}}$ (yrs)	$\mu_m^{\mathtt{act}}$ (yrs)	$t_{\mathtt{act}}$ (hrs)[a]	$\mu_s^{\mathtt{pre}}$ (yrs)	$\mu_m^{\mathtt{pre}}$ (yrs)	$t_{\mathtt{pre}}$ (hrs)[a]	$\frac{\mu_m^{\mathtt{pre}}}{\mu_m^{\mathtt{blk}}}$	$\frac{t_{\mathtt{act}}}{t_{\mathtt{pre}}}$	Error (%)[b]
ibmpg1	6K	11K	709	7.91	11.74	0.001	3.57	6.10	0.08	3.62	6.52	0.03	0.56	2.67x	6.89
ibmpg2	62K	61K	462	1.20	2.94	0.06	4.22	6.63	0.87	4.19	6.79	0.31	2.31	2.81x	2.42
ibmpg3	410K	401K	8.1K	0.90	3.91	0.90	3.35	6.09	7.99	3.35	6.67	4.27	1.71	1.87x	9.52
ibmpg4	475K	465K	9.6K	0.64	2.01	1.10	4.11	8.77	13.73	4.07	9.87	6.81	4.91	2.02x	12.54
ibmpg5	249K	496K	2K	1.49	2.90	0.16	4.17	5.94	0.53	4.18	6.53	0.25	2.25	2.11x	9.96
ibmpg6	404K	798K	10.2K	1.04	1.96	0.33	3.88	8.96	4.64	3.88	9.39	2.07	4.78	2.24x	4.79
ibmpgnew1	316K	698K	19.5K	1.56	3.20	0.17	3.74	12.13	0.47	3.68	13.25	0.42	4.15	1.12x	9.23
ibmpgnew2	718K	698K	19.5K	0.84	4.63	2.99	3.63	5.30	3.21	3.58	5.73	2.60	1.24	1.23x	8.04

[a] $t_{\mathtt{blk}}$, $t_{\mathtt{act}}$ and $t_{\mathtt{pre}}$ denote the run-time(s) for Black's model, Filtering and predictor based approaches, respectively.
[b] The error is calculated for mesh MTF between the *Filtering* and *Filtering+predictor* approaches

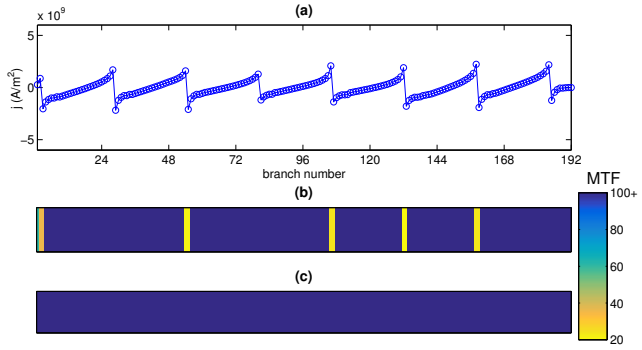

Figure 8: (a) Low current density profile and possible failure locations as predicted using (b) Extended Korhonen's model and (c) Black's model

In order to show the inaccuracy in Black's model, we present two scenarios, based on an interconnect tree from ibmpg2, which is a straight metal stripe with 192 branches. In the first scenario, we apply a high current density profile to the tree (see Fig. 7). In this case, Black's model predicts the first failure time of about 4 yrs, whereas the actual failure time found using the Extended Korhonen's model is around 7.8 yrs. In the second scenario, we apply a low current density profile (see Fig. 8). Here, due to the Blech effect, Black's model predicts that no failure would occur. However, accounting for the material flow between the branches, we found that the first failure would occur around 20 yrs. Thus, Black's' model was pessimistic in the first scenario and optimistic in the second one.

7. CONCLUSION

We proposed a fast physics-based EM checking approach for on-die power grids that removes the unrealistic assumptions inherent in traditional industrial tools. Computational speed is achieved using an efficient filtering scheme and a fast predictor-based approach, with minimal impact on accuracy. The MTFs estimated using our physics-based approach were 3x longer on average than those based on a (calibrated) Black's model. The method is quite fast and is suitable for very large power grids.

8. REFERENCES

[1] I. Bebu and T. Mathew. Confidence intervals for limited moments and truncated moments in normal and lognormal models. *Statistics & Probability Letters*, 79(3):375 – 380, 2009.

[2] J. R. Black. Electromigration-A brief survey and some recent results. *IEEE Trans. on Electronic devices*, 16(4):338–347, Apr. 1969.

[3] S. Chatterjee, M. Fawaz, and F. N. Najm. Redundancy-Aware Electromigration Checking for Mesh Power Grids. In *IEEE/ACM Int. Conf. on Comput. Aided Design*, pages 540–547, San Jose, CA, Nov. 2013.

[4] J. Clement. Electromigration modeling for integrated circuit interconnect reliability analysis. *IEEE Trans. Dev. Mat. Rel*, 1(1):33–42, Mar 2001.

[5] C. L. Gan, C. V. Thompson, K. L. Pey, and W. K. Choi. Experimental characterization and modeling of the reliability of three-terminal dual-damascene Cu interconnect trees. *J. Appl. Phys.*, 94(2):1222–1228, 2003.

[6] S. P. Hau-Riege and C. V. Thompson. Experimental characterization and modeling of the reliability of interconnect trees. *J. Appl. Phys.*, 89(1):601–609, 2001.

[7] X. Huang, Y. Tan, V. Sukharev, and S.-D. Tan. Physics-based Electromigration Assessment for Power Grid Networks. In *ACM/EDAC/IEEE Design Automation Conf.*, pages 1–6, June 2014.

[8] M. A. Korhonen, P. Borgesen, K. N. Tu, and C. Li. Stress evolution due to electromigration in confined metal lines. *J. Appl. Phys.*, 73(8):3790–3799, 1993.

[9] B. Li, J. Gill, C. Christiansen, T. Sullivan, and P. S. McLaughlin. Impact of via-line contact on cu interconnect electromigration performance. In *IEEE Int. Rel. Phys. Symp.*, pages 24–30, April 2005.

[10] D.-A. Li, M. Marek-Sadowska, and S. Nassif. A method for improving power grid resilience to electromigration-caused via failures. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, 23(1):118–130, Jan 2015.

[11] J. Lloyd and J. Kitchin. The electromigration failure distribution: The fine-line case. *J. Appl. Phys.*, 69(4):2117–2127, Feb 1991.

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

[13] S. R. Nassif. Power grid analysis benchmarks. In *ASP-DAC*, pages 376–381, 2008.

[14] W. Schiesser. *Computational Mathematics in Engineering and Applied Science: ODEs, DAEs, and PDEs*. Taylor & Francis, 1993.

[15] V. Sukharev, A. Kteyan, and X. Huang. Postvoiding stress evolution in confined metal lines. *IEEE Transactions on Device and Materials Reliability*, 16(1):50–60, March 2016.

[16] L. Ting, J. May, W. Hunter, and J. McPherson. AC electromigration characterization and modeling of multilayered interconnects. In *IEEE Int. Rel. Phys. Symp.*, pages 311–316, March 1993.

[17] N. Weiss, P. Holmes, and M. Hardy. *A Course in Probability*. Pearson Addison Wesley, 2005.