A Retargetable and Accurate Methodology for Logic-IP-internal Electromigration Assessment

Palkesh Jain[1]
Automotive Systems Group
Qualcomm India, Bangalore
palkesh@qti.qualcomm.com

Sachin S. Sapatnekar
Department of ECE
University of Minnesota
sachin@umn.edu

Jordi Cortadella
Department of Computer Science
Universitat Politècnica de Catalunya
jordi.cortadella@upc.edu

***Abstract*— A new methodology for SoC-level logic-IP-internal EM verification is presented, which provides an on-the-fly retargeting capability for reliability constraints. This flexibility is available at the design verification stage, in the form of allowing arbitrary specifications (of lifetimes, temperatures, voltages and failure rates), as well as interoperability of IPs across foundries. The methodology is characterization- and reuse-based, and naturally incorporates complex effects such as clock gating and variable switching rates at different pins. The benefit from such a framework is demonstrated on a 28nm design, with close SPICE-correlation and verification in a retargeted reliability condition.**

I. Introduction

Electromigration (EM) is a major product aging mechanism revolving around the containment of the average and RMS current densities in interconnects, requiring *cell-external analysis* for signals and power nets connecting cells and *cell-internal analysis* for wires within a logic-IP (standard cell) or a mixed signal IP block. Recently, a great deal of innovation and improvement has been seen on the verification and design strategies for cell-external signal and power grid EM [1]. However, there has not been adequate focus on the robust design and reuse of the standard cells. Ensuring EM reliability for standard cells in a design implies that the exact context at which the cells are used must be bounded to guarantee its robustness in the design. This context could be stated in terms of design limits (loads, slews, frequencies, supply voltage), or reliability (temperature, lifetime, or a failure rate specification tied to current density limits). Without rigorous assessments, a set of logic-IPs designed for foundry A and a *reliability condition* (*e.g.*, 1.2V, 105C, 100k power-on hours (POH), 0.1% cumulative failure and 10C Joule heating (JH) limit) cannot be guaranteed to be EM-safe at another condition (*e.g.*, 1V, 115C, 200kPOH, 0.01% cumulative failure and 15C JH limit) in foundry B. Nevertheless, tradeoffs on these constraints are increasingly in demand in industry due to accelerated inroads of semiconductor houses into newer businesses with different reliability demands [2]. For example, automotive designs, unlike their consumer counterparts, demand more stringent operating conditions than traditional computing applications [3]. From an EM standpoint, meeting this is challenging, as seen from Fig. 1, which highlights the representative current density across various temperature and lifetime specifications. As can be seen, from the mobile to the automotive application domain, the current carrying capability tightens over 20x.

One way to meet such specifications is to approach design in a bottom-up manner with a fresh logic-IP portfolio that meets targeted domain-specific reliability specifications. However, this is prohibitively expensive. A 28nm high-performance library can be used in several applications; even in the same SoC, different blocks (*e.g.*, an ARM core or a DSP) may have different reliability requirements. Due to economic and design effort considerations, product integration over all application domains should be based on the same IP portfolio, with *suitable retargeting* to the application domain.

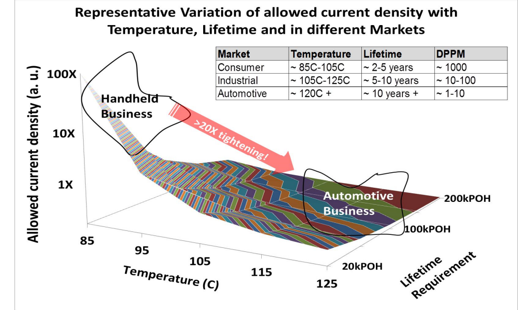

Fig. 1: Typical current density limits as a function of temperature and lifetime, showing >20X differences between consumer and automotive requirements.

A first step towards assuring logic-IP reliability in a design, implemented in production tools [4,5], uses a precharacterized table to model the tradeoffs in various cell parameters (Fig. 2a). Fig. 2b represents a model with two operating constraints: frequency and load (f-L), at fixed values of other parameters. As the operating load increases, the current flow in the cell increases, worsening EM, and the frequency at which it can operate reliably is lower. Such a model can be used at the chip level to determine the safe frequency (f_{safe}) of an instance for any design/reliability condition, and then understand whether an instance can be safely used. Given the compute-intensive nature of characterizing this model (the f_{safe} at each load is computed using a binary search over frequency), the model parameters are limited to frequency, loads and slews, fixing the supply voltages and reliability conditions. This implies that retargeting to other supply voltages or reliability conditions requires fresh (and expensive) recharacterization.

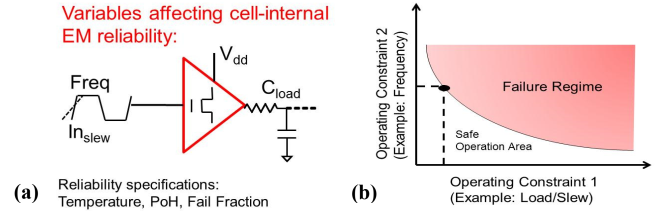

Fig. 2a: Variables affecting cell-internal EM. Fig. 2b: Traditional EM verification using the safe operating region concept.

Our goal is to address some major limitations, associated with such chip-level cell-internal EM analysis:

- **L1 –** Inability to incorporate the impact of arbitrary switching rates on inputs pins and effects such as clock gating in the eventual safe frequency (as illustrated in Fig. 2b): we overcome this by discretely characterizing the individual current components (switching or leakage)

[1] This work was done while the author was with Texas Instruments, India.

- **L2 –** Inability to retarget reliability specifications rapidly for different reliability conditions across markets (Fig. 1), or across block-level reliability specifications in an SoC: we present closed-form formulae for an "equivalent stress" for a cell given its use condition. We refer to this as *on-the-fly retargeting*, based on a specification during design verification, as against a full library characterization at each reliability condition or the compute-intensive task of Fig. 1.

The core methodology of our work naturally enables the model-retargeting by separating out the current computation part versus the verification (unlike the tight coupling in the model of Fig. 2b). We assume the instance operating-frequency (f_{op}) to be an input to our model: derived from PrimeTime [4], VCD, etc., in a typical design flow.

We now look at the basic concepts of EM, along with the details on traditional model in section II. We discuss the proposed model, accounting for arbitrary switching rates and clock gating in section III, followed by reliability retargeting framework and later, results from a production setup of 28nm.

II. EM Modeling – Basic Framework

A. Electromigration Basics

In this section, we review the key parameters affecting EM. In our terminology, we refer to metal segments of the cell as *resistors*. These resistors are obtained by parasitic extraction, which retains key information such as the width, length, and the metal-level for every resistor in the net-list.

Since EM is a statistical process, the time to failure for metal segments stressed in similar conditions also varies. Industrial markets demand low failure rates (*e.g.*, 100 defective parts per million (DPPM) over the lifetime). Chip reliability engineers translate this chip-level metric, to specific fail fraction targets (*FF*), in units of failures-in-time (FIT) on individual resistors.

The classic Black's equation [6] relates the mean time to failure (t_{50}, time to failure for half of the population) to the average current density J across the interconnect cross-section and the wire temperature T as below, where Q is the activation energy, k_B is Boltzmann's constant, n a current exponent (between 1-2), and A is a fitting parameter.

$$t_{50} = A\, J^{-n}\, e^{Q/k_B T} \tag{1}$$

Black's equation predicts the time to failure, and in practice, it is predominantly used to determine the average current density thresholds to meet a target *FF*. It has been demonstrated that *FF* follows a lognormal dependency on the time to failure (t_f, also known as stress time), [6]. The lognormal parameter (z), relates to the time-to-failure as follows, where σ is the standard deviation of the distribution:

$$z = \frac{\ln(t_f) - \ln(t_{50})}{\sigma}\ ;\ FF = \int_{-\infty}^{z} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx \tag{2}$$

Eqs. (1) and (2) are intuitive. For example, for a fixed stress time, t_{50} decreases with increasing stress temperature (eq. (1)), thereby keeping z, and eventually the *FF*, high.

In signal wires, currents flow in both directions, leading to a phenomenon known as AC EM, where the alternating nature of the stress causes limited damage-recovery. This behavior is incorporated by introducing a recovery factor, $\mathfrak{R}$, that is empirically obtained and must be used for adjusting the computed average current in Black's equation [7], as:

$$J = J_{avg}^{+} - \mathfrak{R} J_{avg}^{-} \tag{3}$$

Here, J_{avg}^{+} and J_{avg}^{-} indicate the average current density during current conduction in the positive and negative directions, respectively. Additionally, the wire heating (ΔT) has an inherent dependence on the RMS current, J_{RMS}, as [1]:

$$\Delta T = c * J_{RMS}^{2} \tag{4}$$

Eq. (4), with c as a fitting parameter, follows directly from heat conduction principles. Typically, a limit on the maximum temperature rise due to JH is a design constraint, placing an automatic RMS limit. Pioneering work by Hunter [8] combined the two effects of average EM fails and RMS induced JH in a self-consistent manner, unifying the checking criteria. Given a stress temperature, lifetime, and JH limit constraints, the EM limits for all wires in a cell can be found.

B. Traditional Approach for Modeling EM Reliability

We revisit the traditional approach for verifying logic-IP EM reliability, outlined in Fig. 2b. Given the physical design of the cell, EM verification requires a model that provides a tradeoff amongst various operating conditions such that within the bounds of those tradeoffs, the cell remains EM-clean. The generation of this model requires an iterative search: for example, in Fig. 2b, at a fixed loading and reliability condition (say, 50fF, 1.0V, 105C, 100kPOH), an iterative search over the frequency space is required to determine the maximum f_{safe}, where all resistors within the cell pass EM. This is computationally expensive, since each iteration involves a SPICE-simulation-based verification. A typical optimized procedure requires ten binary search iterations at each loading condition. For a single input cell, whose operating load/slew space is covered through 8x8 matrix in the liberty file, the number of required iterations are about 64x10 = 640 for fixed values of other parameters (supply voltage and reliability specifications). To support operation at multiple voltages, as well as cell reuse across applications, this number must be multiplied by the number of use cases, resulting in a formidable characterization overhead.

While this may even be tractable for single-input cells, for multiple-input cells, this characterization becomes challenging, not just from a compute point of view, but also from the fundamental modeling (**L1**) viewpoint. To illustrate this idea, consider the example of a two input clock-tree mux cell that is used to alternate amongst clocks for downstream propagation. In this experiment, both the input pins (Clk1, Clk2) switch at 100%, but the select pin is toggled to allow passing of first and second clocks in varying amounts (going from 0% to 100% in steps of 25%).

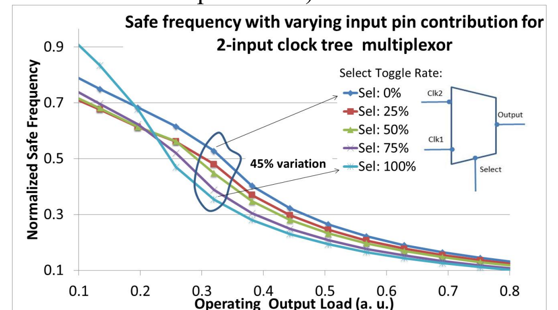

Fig. 3 F_{safe} plot for a 2-input clock-multiplexor cell. Both input clocks switch at 100%, while the select pin chooses one of them, with varying likelihoods.

The f-L plots for the five cases are shown in Fig. 3, and show a variation of up to 45% in f_{safe} estimates, depending on

how often Clk1 or Clk2 is selected over the lifetime, but the traditional model chooses the pessimistic f_{safe} over all cases. Such an asymmetric response can only be captured by the traditional model by individually generating and storing the f-L data for various stimuli, which is expensive in terms of data-generation and storage. Further, effects like clock and power gating are not straightforward to handle in traditional model. Another major drawback (**L2**) is that the traditional model is locked to a particular reliability specification (supply voltage, temperature, lifetime, and failure rate target). Enumerating the model over all reliability specification values is possible but impractical due to its prohibitive cost.

III. Addressing L1 – Incorporating Arbitrary Switching and Clock Gating in Frequency Estimation

A. Library Level Current Characterization

In order to build the model which can help predict the reliability of a cell for arbitrary switching scenarios, we begin in an *ab initio* manner by trying to classify the current flow in the cell as either leakage or switching current. We observe that for a combinational cell with *m* inputs, 2^m distinct static states (various combinations of input pins at logic 1 or 0) are possible. Each of these states can have different leakage flow. Additionally, based on the cell-functionality, there could be several paths (later referred as arcs) from an input pin resulting in an output transition. Every such output transition, causes a switching current flow in the cell-internal resistors.

Thus, first step in our approach is to discretely characterize the current flow through every resistor in the cell, in every legal state (for leakage current) or arc through the cell (for switching current), at all load-slew conditions in the liberty file (typically 8x8). Such a characterization will be used to compute the eventual effective current through any resistor as a weighted summation of the currents in unique scenarios, coupled with the information of arc switching rates and probabilities of legal state occurrences. *The salient feature of our characterization is that it remains independent of the reliability condition: which is actually a chip-level input.* We now talk individually about leakage and switching currents.

A.1 Leakage Current Induced Electromigration

Leakage currents in the cell depend only on the static states of inputs. Therefore, by cycling through all possible input states, we can readily obtain, through SPICE simulations, the current through each resistor (note that average and RMS remain same). For an *m*-input gate, let the leakage current through resistor R_j for a state *k* (of 2^m states) be denoted by $J_{avg,Rj,k}$. Then, the effective leakage through resistor R_j covering all the states and with recovery (eq. (3)) would be:

$$J_{avg,leak,R_j} = \sum_{k=1}^{l} P_k J_{avg,R_{j,k}} - \Re\left(\sum_{i=1}^{2^m-l} P_i J_{avg,R_{j,i}}\right) \quad (5)$$

Here *l* is the number of states with positive current, and P_q is the probability of occurrence of state *q* (a function of the input duty cycles and probabilities).

A.2 Switching/Charging Current Induced Electromigration

Unlike the leakage case, switching currents are tied to a particular arc from a particular input pin of the cell to an output of the cell, through a fixed cell-internal path, with other inputs in noncontrolling states that enable a transition on this arc. For example, for a three-input AOI gate (Y = !(A + BC)), the output Y can fall because of a rise on A in three different states of BC, namely, 00, 01 and 10. Hence, for this particular A→Y arc, the current must be computed through R_j for these three states of BC. We can leverage the simulation framework of industrial timing characterization systems, [9], to obtain information about all such arcs and states through the cell. For a particular arc *i* and associated non-controlling state *k*, we denote the time duration over which this current is calculated as s_{ik}. A similar convention is followed by $R_{j,ik}$. Then, the effective positive current through R_j ($J_{avg,sw,Rj}$) is given by:

$$J_{avg,sw,R_j} = \sum_{i=1}^{all\ arcs}\left(\sum_{k=1}^{all\ states} N_{ik} J_{avg,R_{j,ik}} \frac{s_{ik}}{P}\right) \quad (6)$$

Here, N_{ik} and *P* are the design-level variables – respectively, the number of transitions a particular arc (and associated states) sees in the lifetime, and the switching period. Similar calculations for RMS currents ($J_{rms,sw,Rj}$) yield:

$$J_{rms,sw,R_j} = \sqrt{\sum_{i=1}^{all\ arcs}\left(\sum_{k=1}^{all\ states} N_{ik} J^2_{rms,R_{j,ik}} \frac{s_{ik}}{P}\right)} \quad (7)$$

Since we leverage the timing characterization framework, we do not compute s_{ik}, instead reuse it from timing setup, [9]. Also, s_{ik} is typically greater than the delay itself, and therefore accurately captures the tail effects.

B. Effective Current Estimation for a Chip-Level Instance

After extracting the leakage and switching currents, we now present an example of the calculations for a simple single-input clock tree cell. Denoting the *chip-level inputs:* activity-rate-adjusted frequency by *f*, and the input duty cycle (signal high probability) by *α*, we can arrive at the average current ($J_{avg,Rj}$) for any resistor R_j in the cell, using eqs. (6), (7), as:

$$J_{avg,R_j} = f\left(J_{avg,R_j,rise} s_{rise} - \Re J_{avg,R_j,fall} s_{fall}\right) + \left((1-\alpha)L_{0,R_j} - \Re\alpha L_{1,R_j}\right) \quad (8)$$

$$J_{rms,R_j} = \left[f\left(J^2_{rms,R_j,rise} s_{rise} + J^2_{rms,R_j,fall} s_{fall}\right) + (1-\alpha)L^2_{0,R_j} + \alpha L^2_{1,R_j}\right]^{1/2} \quad (9)$$

Here, $J_{avg,Rj,rise}$ and $J_{avg,Rj,fall}$ represent the average currents through the resistor R_j during the rise and the fall operations, individually. The currents $L_{0,Rj}$ and $L_{1,Rj}$ represent the leakage through the resistor R_j in the input logic low and high states respective. Additionally, s_{rise} and s_{fall} denote the current calculation duration for the rise and fall arcs, respectively. Similar notations hold true for the RMS currents in eq. (9). Indeed, eqs. (8), (9) are capable of estimating current through the resistors for any chip-level inputs: *f* and *α* .

Next, we look at incorporating clock-gating in the formulations. We notice that as a phenomenon, clock gating can occur in an arbitrary way over the lifetime of the chip. For instance, the clock could be gated for a fixed number of cycles, periodically. Such uniform gating is akin to a direct reduction in the operating frequency; readily approximated by specifying the activity-rate adjusted frequency in above eqs. (8), (9). However, the cases when the clock gating is nonuniform, or is uniform only in the intervals, are nontrivial. The key determinant here is the thermo-mechanical time constant of JH in interconnect (typically in few milliseconds for copper), which signifies the duration for interconnect to respond to the RMS current. Hence, if the time interval between successive clock gating events is larger, then, the full

current (without activity correction) should be used for reliability estimations with appropriate durations.

We defer treatment of nonuniform clock-gating as a future work (although our framework is capable), and focus on the uniform case. This makes the solution similar to setting a pin specific activity rate. Hence, for a 1GHz clock-tree element, gated-high for 25% of the lifetime, we would input corrected f as 750MHz in eq. (8), (9), and α as 0.375 (assuming 50% clock duty cycle). Thus, the computation procedure can be:

Algorithm 1 Current computation through every resistor

Input: Timing characterization setup; **Output:** $J_{avg,Rj}$ and $J_{rms,Rj}$
1. **for each** cell in the library; **for** every load/slew in (8x8 matrix)
2. simulate for every legal input-state combination
3. **for each** resistor R_j, store $J_{avg,Rj,k}$ **end**
4. simulate for every legal switching scenario (arc)
5. **for each** resistor R_j, store $J_{avg,Rj,ik}$ and $J_{rms,Rj,ik}$ **end**
6. **end** cell characterization
7. **for each** instance in the design
8. estimate f, α, slew for all the input pins and loading
9. **for each** resistor R_j of the instance at chip level
10. **query-and-add** various current-components, eqs. (8), (9)
11. **end;** store $J_{avg,Rj}$, $J_{rms,Rj}$ at given conditions
12. **end**

It must be mentioned here that RMS formulations work under the assumption that the different components of the current (leakage and switching) are non-overlapping. However, this only leads to marginal errors, as we will see later in section V.

C. Instance-Safe-Frequency Estimation at Chip Level

Once we have estimated the currents in the cell-internal segments, the EM checking procedure at chip-level can subsequently be approached in two manners:

- **Predict the safety of the cell** (pass or fail), given a full set of operating conditions of the cell.
- **Calculate a safe operating parameter** of the cell, provided the rest of the operating conditions, that efficiently checks the cell-EM-safety under the specified operating conditions.

The first is rather trivial, since eqs. (8), (9) and Algorithm 1 lend themselves readily to allow substitution of the exact operating conditions, and subsequent verification of currents (through all resistors) against the foundry EM thresholds.

In real designs, however, the actual operating frequency of the instance may depend on the usage across a wide variety of applications. A recommendation of maximum f_{safe} is therefore necessary, by working the problem backwards. Such estimation requires careful derivation from a self-consistent viewpoint. Our approach here provides closed-form solutions for the f_{safe} that do not require an intensive SPICE-simulation-based binary search, as in traditional methods.

Let $J_{avg,th}(T,t)$ and $J_{rms,th}(\Delta T)$ represent the average and RMS current density limits, as a function of stress temperature, lifetime, and maximum heating constraint. From eqs. (8), (9), by setting the left-hand sides to the threshold densities, we can actually constrain the RMS or average-limited frequencies ($f_{max,RMS,Rj}$ and $f_{max,AVG,Rj}$, respectively) in following manner:

$$f_{max,RMS,R_j} = \frac{J^2_{rms,th}(\Delta T) - \left((1-\alpha)L^2_{0,R_j} + \alpha L^2_{1,R_j}\right)}{\left(s_{rise}J^2_{rms,R_j,rise} + s_{fall}J^2_{rms,R_j,fall}\right)} \tag{10}$$

$$f_{max,AVG,R_j} = \frac{J_{avg,th}(T,t) - \left((1-\alpha)L_{0,R_j} - \Re\alpha L_{1,R_j}\right)}{\left(s_{rise}J_{avg,R_j,rise} - \Re J_{avg,R_j,fall}s_{fall}\right)} \tag{11}$$

We can now apply the self-consistent formulations, [8], to estimate the safe parameter (frequency) of the resistor. The entire process has to be approached iteratively, as shown in Algorithm 2, to determine the f_{safe} for an instance, which can be then used as a design constraint.

Algorithm 2 Self-consistent safe-frequency estimation of the cell

Input: $J_{avg,Rj}$ and $J_{rms,Rj}$ from Alg. 1; **Output:** $f_{max.inst}$ of the cell
1. **for every** instance of the design
2. **start** with an assumption on $f_{max.inst}$
3. **for every** resistor R_j of the instance
4. start with an estimate of $f_{max.RMS,Rj}$ (from eq. (10))
5. estimate ΔT; **if** ΔT > limit; **break**
6. estimate $J_{avg,th}(T+\Delta T)$
7. estimate $f_{max.AVG,Rj}$, using $J_{avg,th}(T+\Delta T)$ (from eq. (11))
8. **iterate** till $f_{maxAVG,Rj}$ and $f_{maxRMS,Rj}$ are close
9. **if** ($f_{maxAVG,Rj} < f_{max.inst}$); $f_{max.inst} = f_{maxAVG,Rj}$
10. **end;** return $f_{max.inst}$
11. **end**

To evaluate this procedure, we revisit the two-input clock tree mux from the earlier discussion around Fig. 3. Fig. 4 provides the f_{safe} plot for this case, for a fixed operating condition and output load, showing the results of binary-search-based SPICE simulation, our approach, and the traditional method that chooses the f_{safe} pessimistically over all switching conditions. We see that the proposed model fits the SPICE behavior very well, modeling the arbitrary switching rates on different pins, as against the large pessimism in the traditional approach. For all library cells, we obtain similar accuracy, but data for the whole library is not shown here due to space limitations.

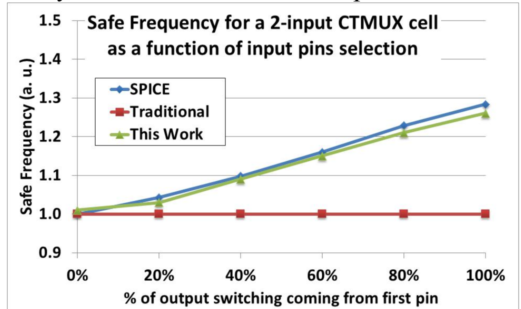

Fig. 4. F_{safe} for the circuit in Fig. 3, at selected load/slew. F_{safe} varies based on the extent of switching coming from first or second pin. The proposed model captures the behavior fully, unlike the excessively pessimistic traditional one.

IV. Addressing L2 – On-the-Fly Retargeting of Reliability for Arbitrary Specifications

The formulations of previous sections were all dependent on the library data, characterized at one set of reliability conditions. However, as described in Section I, there is an increasing need for *on-the-fly* reliability retargeting at verification stage, as the IP library gets used under different conditions. As noted earlier, meeting such a goal is impractical traditionally, as a full characterization of the entire IP library (Fig. 2b) is required for this at each new condition.

Our core methodology enables the ability to perform this retargeting efficiently, since the characterization is disjoint from the reliability condition (whereas these are tightly coupled in the traditional methodology). We begin by revisiting eq. (1), (2), relating EM lifetime and lognormal z:

$$\sigma z = \ln(t_f) - \ln(A) + n\ln(J) - \frac{Q}{k_B T} \tag{12}$$

Thus, the reliability of two different sets of stresses, denoted by subscripts a, b, can be related as follows (governed by same fitting parameter A, but other terms in eq. (12) differ):

$$\sigma z_{cond,b} = \sigma z_{cond,a} + \left[\frac{ln\left(\frac{t_b J_b^n}{t_a J_a^n}\right) -}{\frac{Q}{k_B}\left(\frac{1}{T_b+\Delta T_b} - \frac{1}{T_a+\Delta T_a}\right)} \right] \quad (13)$$

Here, the variables t, J, T, and ΔT represent the stress time, current densities, temperature and JH, respectively, while the subscripts a and b refer to the two different conditions.

This equation is a powerful representation of the scaling factors that must be used to assess the equivalence. For example, to find the equivalent temperature that causes the same reliability loss as the stress of increased lifetime, we must set $z_{cond,b} = z_{cond,a}$ and obtain the T_b from eq. (13) as a function of (T_a, t_a, t_b, J_a, and J_b). The basic idea here is that EM aging is either accelerated or decelerated by the change in operating conditions. Hence, if the circuit is aged under a new condition b, then the equations here help *transpose* the stress to the known (characterized) condition a, with one of the stress parameter of condition a, and may become either more severe or benign than condition b. Next, we use eq. (13) for on-the-fly reliability retargeting for several cases:

Case I: Variations in Temperature If the use temperature and/or POH specification are different from the original conditions, then it is straightforward to address this by using eq. (13) to determine new current density thresholds, and then updating the f_{safe} in eq. (11). Such a modification only affects the average, and not the RMS reliability, which depends only on the ΔT constraint for the RMS rule: if this changes, we can update $J_{rms,th}(\Delta T)$ and then eq. (10).

A second situation is the common industry scenario when the stress profile is provided by the user as a time profile, as the series $\{(J_1,T_1,t_1), (J_2,T_2,t_2), \ldots, (J_m,T_m,t_m)\}$, *i.e.*, from time t_{k-1} to time t_k, it experiences current stress J_k at temperature T_k. If the baseline stress is characterized at (J_0, T_0), then we can relate the kth stress vector to the baseline stress at (J_0,T_0) with an equivalent stress time $t_{k,0}$. In other words, the stress at temperature T_k is transposed to an equivalent stress time at temperature T_0. Consequently, our stress retargeting scheme maps the entire stress to (J_0,T_0,$t_{eq,0}$), where $t_{eq,0} = \sum_{k=1}^{m} t_{k,0}$.

Case II: Variation in Operating Voltage If the eventual use voltage of the library is different from the characterization voltage, current scaling must be performed. Such a scaling is straightforward in our framework, since the leakage and switching related components are separately stored, as described in eqs. (6), (7). Based on our experiments, we see that a linear scaling works very well for voltage, while an exponential model is required for leakage. A second situation (arising due to power-management), is when the voltage is represented as a series: $\{(V_1,t_1) \ldots ((V_m,t_m)\}$. In such a case, we can follow the scaling procedure to obtain a current series, which can then be dealt like Case I.

Case III: Variation in Failure Rate Specification The $J_{avg,th}(T,t)$ in eq. (11) is really a function of the fail fraction FF, which in turn is a function of z (eq. (2)). Therefore, z_{target} is the inverse function of FF_{target}. Hence, if the FF specified by the end user changes from, say, 0.1% to say 0.01% cumulative, it can be readily translated to z, translated to a current density limit using eq. (13), and then used in eq. (11) for verification.

Fig. 5 shows a graphical representation of such an on-the-fly retargeting using the proposed model from a representative cell. For ease of exposition, we use a fixed slew, as in Fig. 2b. As before, we show results from a representative cell due to space constraints. Other library cells show similar trends.

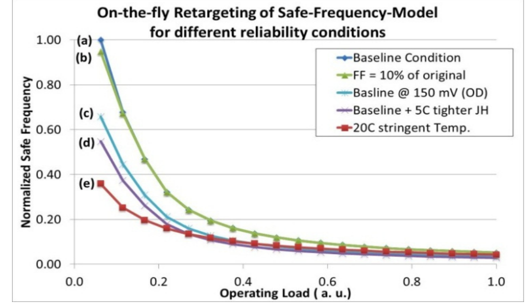

Fig. 5 Demonstrating on-the-fly retargeting of the basic frequency-load curve (Fig. 2b) with changes in the constraining criteria (at a fixed slew point)

Curve (a) represents the reliability at the baseline condition. If the FF requirement of the design changes and drops to 10% of the original, the curve slides down to (b) due to reduction in EM capability at tighter FF requirement. The drop is not drastic as this specific cell is RMS-current-limited, rather than being limited by the average current. Similarly, if the use-voltage has a 150mV overdrive over the characterized value, the reliability is represented by curve (c), which shows degraded reliability due to increased current flow. Similar behavior is seen in curve (d) if the JH (RMS current specification) is tightened by 5C. Finally, a 20C higher temperature requirement is modeled by curve (e) – almost 3X tightening! We reiterate that while our approach naturally handles above case studies, traditionally, they would require a complete f_{safe} recharacterization at various conditions.

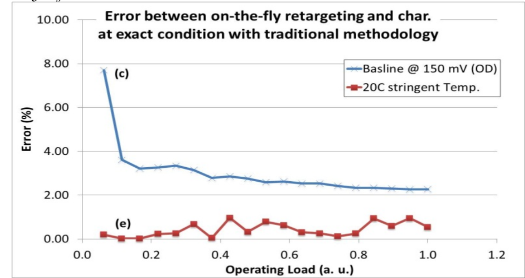

Fig. 6 Validation of retargeting methodology versus SPICE for two conditions, (c) and (e), of Fig. 5.

We validate this retargeting by direct comparison in curves of Fig. 5 to those of traditional methodology (obtained by the actual characterization at the *exact* condition). We show the percentage error for two conditions, (c) and (e) in Fig. 6. For (e), where the temperature specification is altered, the required retargeting only affects the verification part (as the current density limits are scaled), which incurs little error. For (c), the retargeting is due to 150mV overdrive, where we use a more approximate current-scaling model. The error here is acceptable, particularly considering the fact that the higher errors of over 5% all correspond to a lower load regime, which is normally a low-current, EM-safe zone.

V. Production Design Analysis

We now examine the application of our methodology in an industrial scenario, discussing the setup and the workflow. For the production design analysis, we took a high performance, 28nm block (2mmx2mm; ~600K instances, >10M transistors),

operating at 1GHz clock frequency. The block is part of a large industry SoC. We compare the characterization as well as the final data application, with entire flow being outlined in Fig. 7 for the proposed method. The new method, in essence, is a three-step process: (a) IP library characterization at a baseline condition, (b) determining the reliability constraints for this design and (c) integration into the timing/ implementation tool. The true retargeting flexibility of the proposed approach comes in form of (b), which is an on-the-fly input to verification completely detached from (a). The flow of (c) uses a standard industrial design methodology.

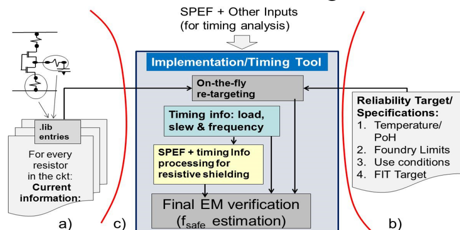

Fig. 7. Overall methodology and data-flow diagram for the proposed method

A. Library Characterization

The entire library of a few thousand cells was characterized in two ways: (a) a full SPICE-based approach for generation of the traditional f_{safe} model at a fixed baseline reliability condition, and (b) the proposed methodology. Using parallelized and multi-threaded SPICE simulations (Cadence Spectre), the runtime for (a) was about 800 CPU hrs, while 80 CPU hrs for (b). For (b), the arcs and states information from timing characterization framework [9] is reused.

B. Final Reliability Verification

The final application of the library-generated data was performed in the timing tool (Encounter, [5]), through a custom developed scriptware, which reads in both the characterization data types. The timing analysis was performed at the baseline condition, to arrive at the slews and activity rates through all the input pins. In the traditional approach, we step through every instance in the design and compare the queried f_{safe} to the f_{op}. Note that since this approach suffers from the problems discussed earlier (specifically, L1), a final full-SPICE simulation (with the RC-loading of the driver instance) is required after the initial results from the frequency comparisons. A total of about 600K instances were analyzed in this way, and finally, over 4500 instances with the frequency ratio > 1, were simulated further. The simulation-stimuli were a simple 1010 transition (at f_{op}), since all the instances were single input clock tree inverters, buffers and gaters (only eight unique cells). The final set of violations after the full SPICE simulations came down to 426.

On the other hand, in the new approach we additionally implement Algorithms 1 and 2, and based on the chip level specifications (lifetime/temperature/voltage/fail fraction), the equations are updated on-the-fly for the f_{safe} comparison of every instance. It must be noted that for instances driving a highly resistive network, lumped load usage for querying cell-internal currents becomes pessimistic, [10]. We work around this inaccuracy by developing a custom method to additionally read in the driver parasitics, and de-rate the queried currents.

Finally, we plot the population distribution of frequency ratios in Fig. 8, for three cases: a) traditional, b) proposed and c) proposed method at a retargeted condition. For every method, we plot the ratio of f_{op} to f_{safe}, which signifies the EM criticality for that instance. Hence, an instance with f_{op} greater than f_{safe} (red region in the plot) is deemed as EM failure and must be acted upon for fixing (either by load reduction or replacement). The y-axis shows the distribution of number of instances in design with a particular f_{op}/f_{safe} ratio.

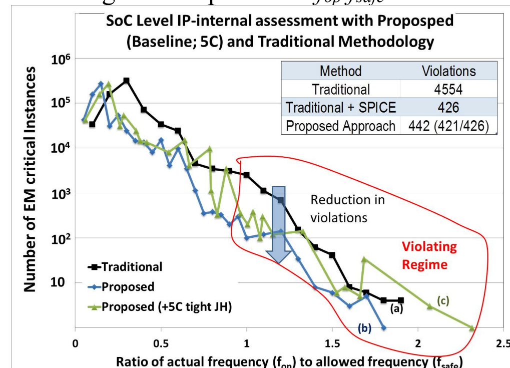

Fig. 8. Distribution plot for a 28nm block (>600K instances), highlighting the number of EM-critical instances for a) traditional, b) proposed and c), proposed method with 5C lower JH. Inset shows reduction in violations.

As we see from Fig. 8, the proposed approach reports a total of 442 violations, 421 of which overlaps with the traditional methodology (+ SPICE). The remaining: false (21) and escaped (5) violations from the new approach were found to be relatively less critical, with frequency ratios in the range of 1.14 to 0.9. Thus, the new approach agrees well with SPICE.

The final retargetability of the proposed approach is evident by the curve c) in Fig. 8, which is run at an additionally tight constraint of 5C lower JH from the baseline. Since a 5C lower JH significantly reduces the allowed RMS currents, we see the number of violations to increase almost by 3X to 1297! Clearly, without a recharacterization, such verification is not possible through the traditional approach.

VI. Conclusion

An accurate and retargetable methodology for IP-internal EM verification was presented in this work. The methodology was shown to be highly flexible, allowing on-the-fly retargeting of reliability at the design verification stage. We presented formulations to incorporate arbitrary switching and clock gating on various pins of the IP. We shared results from a 28nm production setup, demonstrating significant relaxation in terms of violations, close correlation to SPICE and analyses at retargeted reliability conditions, which are plausibly impractical in traditional scheme of things.

References

[1] J. Lienig, "EM impact on physical design in future nodes," *ISPD*, 2013.
[2] C.Yuan *et. al.,* "Optimizing standard cell design for quality," SPIE Advanced Lithography. International Society for Optics and Photonics, 2014.
[3] Q100 specification, http://www.aecouncil.com/AECDocuments.html
[4] Synopsys ICC user guide, http://www.synopsys.com
[5] Encounter Design Implementation tool, (http://www.cadence.com).
[6] J. Black, "Electromigration failure modes in aluminum metallization for semiconductor devices." *Proceedings of the IEEE*, 57:9 (1969): 1587-1594.
[7] K. Lee, "EM recovery under bipolar-and unipolar current," *IRPS*, 2012.
[8] W. R. Hunter, "Self-consistent solutions for allowed interconnect current density. II. Application to design guidelines," *IEEE TED,* 1997.
[9] Altos Library Characterization tool, (http://www.cadence.com).
[10] P. Jain and A. Jain, "Accurate current estimation for interconnect reliability analysis," *IEEE TVLSI*, 2012.