Potential Critical Path Selection based on a Time-Varying Statistical Timing Analysis Framework

P. R. Chithira, *Student Member, IEEE* and Vinita Vasudevan, *Member, IEEE*

***Abstract*—Negative bias temperature instability along with the presence of process variations has resulted in time-varying path criticalities. To ensure reliable circuit operation, aging sensors are used at the end of potential critical paths (PCPs) for delay monitoring. Optimization of the number of delay sensors requires accurate computational models for prediction of criticality and selection of PCPs. We identify a path as a PCP if its maximum global criticality over the lifetime exceeds a certain threshold. However, the global criticality of a path could vary non-monotonically over the lifetime of the device. In this paper, we propose a framework for time-varying SSTA, wherein the circuit delay is obtained as a collection of time-varying canonicals with break points in time which define the end of validity of one and the start of the next canonical. We show that the global criticality of any path will be maximum either at $t = 0$ or at these break points. Hence, criticality computation and PCP selection needs to be done only at these time points, which typically is less than four. The time-varying SSTA is integrated with a previously proposed criticality computation technique to identify the PCPs and the results are validated against Monte Carlo simulations.**

***Index Terms*—Aging, NBTI, statistical timing, path criticality, potential critical path**

I. Introduction

With technology scaling, circuit aging has become a major reliability concern in digital integrated circuits. Aging effects cause transistor parameters to degrade over time, significantly affecting the circuit performance. One of the dominant aging effects is Negative Bias Temperature Instability (NBTI). NBTI degrades threshold voltage of PMOS transistors in inversion region, leading to a progressive increase in circuit delay. Consequently, a circuit that meets timing at testing phase may fail beyond a certain operation time, thus limiting the lifetime of the device. In the presence of process variations, the gate delay distributions change over time depending on the stress factors. This can result in different paths becoming critical not just in different dies, but also at different time points, leading to a time varying path criticality. Any path that can become critical during the expected lifetime of the circuit is referred to as a 'potential critical path' (PCP).

The traditional method to deal with circuit aging is guard-banding. However, this is highly pessimistic as it assumes worst case stress conditions for all the gates and results in significant performance loss. Recently, aging sensors have been proposed for on-line circuit delay prediction in the post-silicon phase [1], [2]. The aging sensors are typically placed at the end of PCPs for delay tracking. Adaptive techniques like frequency scaling, voltage scaling and adaptive body bias are then employed to mitigate the effects of aging [3]. Owing to the overhead associated with insertion of aging sensors, it is infeasible to track the delay of a large number of paths. Also, missing a PCP could result in circuit failure. Therefore, it is essential to have a path selection technique that identifies the optimal set of PCPs necessary for circuit delay tracking.

Several methods have been proposed for the selection of PCPs under NBTI for delay monitoring. The first category of techniques follow a two step procedure. The first step is generation of an initial path set. The potential critical paths are then selected from this initial set using certain path selection criteria. In [4]–[6], both these steps are performed based on STA. The technique proposed in [4] performs a traditional STA at $t = 0$ and enumerates all paths with worst case delay (assuming 20% degradation) greater than the delay of the longest critical path (LCP). A path from this initial set is identified as critical if its aged delay, obtained using an NBTI aware timing library, exceeds the fresh delay of LCP with a timing margin. The method proposed in [5] is similar to that of [4], with the difference that the path delays are compared with respect to a clock period rather than delay of LCP obtained from STA. The approach proposed in [6] is based on timing graph reduction and does not have a separate path filtering step. An edge in the graph is removed if all the paths passing through it is non-critical even under worst case aging conditions. All the paths passing through the remaining edges are then chosen for testing under aging.

The techniques proposed in [7]–[9] use the results from STA for the generation of initial path set. However, the selection of PCPs is done incorporating the effect of process variations. In [7], the mean and standard deviation of fresh (without aging) and aged delays are obtained for each path in the initial set. A path is chosen as a PCP if the probability that its aged delay distribution is greater than the fresh delay distribution of longest critical path (LCP) from STA exceeds a certain threshold. The authors in [8], [9] claim that the local criticality of any path will be maximum at the end points, at $t = 0$ or at the end of lifetime. For every pair of paths p_i and p_j in the initial set, the local criticality of p_i with respect to p_j is computed at the end points and p_i is removed if the maximum criticality is below the threshold. However, this is obvious only

Manuscript received August 31, 2018; revised November 24, 2018; accepted January 7,2019.

The authors are with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai 600036, India (e-mail: ee13d027@ee.iitm.ac.in, vinita@ee.iitm.ac.in).

if the variation in standard deviation with time is negligible.

Another potential critical path selection technique incorporating process variations and based on graph reduction was proposed in [10]. Initially, all edges with low probability of being a part of any potential critical path are removed. The remaining paths in the reduced timing graph are then enumerated. A path in this set is selected as a PCP if the probability that its maximal aged delay distribution is greater than the minimal aged delay distribution of longest path from STA exceeds a given threshold. However, identification of the edges to be pruned is not simple if there are large number of uncorrelated paths with similar delays.

All these techniques require an initial path set and it is not clear how many paths should be selected. To ensure that no critical paths are missed, large number of paths should be enumerated which is time-consuming. Moreover, PCP selection using STA will result in large number of highly correlated paths so that monitoring any one of these paths is sufficient for delay tracking. Even though the path selection using local criticality takes into account the correlations, the local criticality is a good approximation only if the number of paths in the initial set is large. The accurate measure for path criticality is its global criticality [11]. Therefore, an alternative technique for potential critical path selection is to perform block based SSTA at several time instants over the lifetime of the circuit using aged device models [12]–[14]. For path selection, this is followed by criticality computation at these time points. However, the global criticality of paths could vary non-monotonically over the lifetime of the device. This makes it difficult to figure out the time instants at which path selection and criticality computations should be performed. Moreover, existing statistical aged gate delay models have some limitations which are explained in more detail in Section II. In this paper, we address these issues. The main contributions of our paper can be summarized as follows.

1) We derive a simple time-varying canonical delay model that considers the correlation between process variations and NBTI. It also includes the intrinsic charge fluctuations, which are independent of process variations.
2) We develop a time-varying SSTA framework to evaluate circuit delay as a collection of time-varying canonicals, with break points in time. A break point defines the time of switching from one canonical model to the other.
3) We show that the global criticality of any path will be maximum either at t = 0 or at the break points in the circuit delay distribution. Therefore, criticality computation needs to be done only at these time points.
4) We make use of the hierarchical partitioning technique proposed in [15] at these time points to identify the PCPs with global criticality above a specified threshold.

The rest of this paper is organized as follows. In section II, we develop a time-varying canonical model for the representation of gate delays under NBTI. Section III discusses the proposed technique for time-varying SSTA and the implementation details. In Section IV, we derive the properties of time-varying path criticality. The proposed algorithm for potential critical path selection is discussed in Section V. Section VI contains a description of the overall framework. Section VII presents the results and Section VIII concludes the paper.

II. Time-varying Canonical Delay Model

A. Threshold voltage degradation

The threshold voltage degradation due to NBTI at high frequencies is modelled as [16],

$$V_{th,nbti}(t) = A_{nbti}(V_{th}(t=0))\alpha_s^n t^n \tag{1}$$

where α_s equals $\alpha/(1-\alpha)$, α is the stress factor indicating the fraction of time the transistor is under stress and $n = 1/6$. The coefficient A_{nbti} depends on the operating conditions and initial threshold voltage at time $t = 0$ where,

$$V_{th}(t=0) = V_{th0} + V_{th,PV} \tag{2}$$

Here, V_{th0} is the nominal threshold voltage and $V_{th,PV} = \mathcal{N}(0, \sigma^2_{V_{th,PV}})$ is a zero mean random variable representing the process variations in threshold voltage. The dependence of A_{nbti} on process variations is modelled as [17],

$$A_{nbti}(V_{th}(t=0)) = A_0(1 + S_A V_{th,PV}) \tag{3}$$

where $A_0 = A_{nbti}(V_{th0})$ and S_A indicates the sensitivity of A_{nbti} to V_{th} variations normalized with respect to A_0. Clearly, the two random variables, $V_{th,nbti}(t)$ and $V_{th,PV}$, are correlated.

In addition to this, there is an independent intrinsic component within NBTI due to the random charge fluctuations of the generated interface traps. The variance of the intrinsic component is modelled as [18],

$$\sigma^2_{int} = K\mu\{V_{th,nbti}(t)\} \tag{4}$$

where $\mu\{V_{th,nbti}(t)\}$ is the mean of threshold voltage degradation, $K = qt_{ox}/(\epsilon_{ox}A_G)$ and A_G is the transistor area.

The statistical aged gate delay models in [12], [17], [19] considered the correlations between process variations and NBTI, but neglected the variations due to the intrinsic charge fluctuations. The models proposed in [13], [18] and [20] included the random charge fluctuations within NBTI. However, the threshold voltage degradation was modelled as a function of the nominal threshold voltage. The method proposed in [21] used Monte Carlo based transistor simulations to study the effect of process variations on aging and hence, is not suited for the analysis of large circuits. The approach in [14] considered run time voltage and temperature variations and modelled BTI as a log normal distribution. Although comprehensive, it is not in the standard SSTA format since the gate delay was modelled as a mix of normal and log normal distributions.

Considering this, we first derive a simple time-varying model for threshold voltage that takes into account the correlation with process variations as well as the intrinsic fluctuations within NBTI and can very easily be incorporated into the canonical delay model used in SSTA. Using iterated expectations and law of total variance, the mean and variance of

threshold voltage degradation due to NBTI is computed as follows.

$$\begin{aligned}\mu\{V_{th,nbti}(t)\} &= \mathrm{E}_{V_{th,PV}}\{\mathrm{E}\{V_{th,nbti}(t)|V_{th,PV}\}\} \\ &= A_0\alpha_s^n t^n \end{aligned} \quad (5)$$

$$\begin{aligned}\sigma^2\{V_{th,nbti}(t)\} &= \mathrm{Var}_{V_{th,PV}}\{\mathrm{E}\{V_{th,nbti}(t)|V_{th,PV}\}\}+ \\ &\quad \mathrm{E}_{V_{th,PV}}\{\mathrm{Var}\{V_{th,nbti}(t)|V_{th,PV}\}\} \\ &= (A_0 S_A \alpha_s^n t^n)^2 \sigma^2_{V_{th,PV}} + K A_0 \alpha_s^n t^n \end{aligned} \quad (6)$$

Let $\Delta V_{th}(t)$ denote the overall deviation in threshold voltage from the nominal value. Hence,

$$\mu\{\Delta V_{th}(t)\} = \mathrm{E}\{V_{th,PV} + V_{th,nbti}(t)\} = A_0\alpha_s^n t^n \quad (7)$$

Since $V_{th,PV}$ and $V_{th,nbti}(t)$ are correlated, the variance of $\Delta V_{th}(t)$ can be derived as follows.

$$\begin{aligned}\sigma^2\{\Delta V_{th}(t)\} &= \sigma^2_{V_{th,PV}} + \sigma^2_{V_{th,nbti}(t)} + 2\mathrm{Cov}\{V_{th,PV}V_{th,nbti}\} \\ &= (1 + A_0 S_A \alpha_s^n t^n)^2 \sigma^2_{V_{th,PV}} + K A_0 \alpha_s^n t^n \end{aligned} \quad (8)$$

where $\mathrm{Cov}\{V_{th,PV}V_{th,nbti}\}$ is the covariance between process variations and NBTI induced threshold voltage degradation.

In the canonical delay model used in statistical timing analysis, $V_{th,PV}$ is expressed in terms of principal components for modelling the spatial correlations. Therefore, we model $\Delta V_{th}(t)$ in terms of the principal components (X_i) as,

$$\Delta V_{th}(t) = A_0\alpha_s^n t^n + (1 + A_0 S_A \alpha_s^n t^n)\sum_{i=1}^{N} S_{X,i}X_i + S_{int}t^{\frac{n}{2}}R_{int} \quad (9)$$

where $S_{int} = \sqrt{K A_0 \alpha_s^n}$ and R_{int} is an independent component that models the random charge fluctuations in NBTI.

As indicated by (1), the threshold voltage degradation due to NBTI is proportional to coefficient A_{nbti} which in turn depends on the operating conditions. Since a higher supply voltage and temperature result in higher rate of V_{th} degradation [22], a worst case operating condition of ($V_{dd} = 1.32V$, $T = 125^0C$) and ($V_{dd} = 1.1V$, $T = 125^0C$) for 90nm and 45nm technology respectively is considered for our analysis.

The threshold voltage degradation under NBTI also depends on the stress factor (α) and hence is different for different workload conditions at the primary inputs. In this work, we ignore any variations in input workload with time. As the input workload is unknown, a static probability (SP) of 0.5 is assumed at the primary inputs. Our simulation results show that the circuit delay degradation under various input static probability conditions is typically within 5% of the $SP = 0.5$ assumption. The static probabilities as well as the pairwise signal correlations at the gate inputs are calculated from the Value Change Dump (VCD) file after performing logic simulation. The stress factors are then computed using the technique proposed in [23].

B. Gate delay model

Ideally, for an n-input gate, the delay degradation of any input-output arc depends on the threshold voltage degradation of all the PMOS transistors. But, similar to the existing statistical timing analysis techniques under NBTI [14], [17], we assume that the delay degradation of arc (x_i, y) depends only on the threshold voltage degradation of transistor M_i connected to x_i. Using first order Taylor series expansion of alpha power law model, the dependence of gate delay on time-varying V_{th} variations is modelled as [17],

$$d(t) = \mu_{d0} + S_V \Delta V_{th}(t) \quad (10)$$

where μ_{d0} is the nominal delay and S_V is the sensitivity of gate delay to changes in threshold voltage. Therefore, the time-varying canonical delay model can be obtained from (9) as,

$$\begin{aligned} d(t) = \mu_{d0} + d_t t^n + \sum_{i=1}^{N}(S_{d_{X,i0}} + d_{X,i}t^n)X_i \\ + \sum_{j=1}^{M}(S_{d_{R,j0}} + d_{R,j}t^n)R_j + d_{R,int}t^{\frac{n}{2}}R_{int} \end{aligned} \quad (11)$$

Here, μ_{d0} denotes the nominal gate delay at $t = 0$. $S_{d_{X,i0}}$ and $S_{d_{R,j0}}$ denote the sensitivities associated with process variations (X_i) and random components (R_j) at $t = 0$. d_t, $d_{X,i}$ and $d_{R,j}$ denote the sensitivities associated with time dependence which in turn depend on the parameters S_V, S_A, A_0 and α. $d_{R,int}t^{\frac{n}{2}}$ denotes the sensitivity associated with the intrinsic component R_{int}. In comparison with the extended canonical form defined in [24], there are two main differences namely, (a) the mean and sensitivities are functions of time and are of the form $x_0 + x_t t^n$ (b) there is an additional random component (R_{int}) representing the intrinsic fluctuations in NBTI. Note that only rise delays degrade due to NBTI and falling gate delays are independent of time.

III. Time-varying SSTA

A block based time-varying SSTA (TV-SSTA) is used to find arrival time at each node as a function of time. The following subsections describe SUM and MAX operations on delays expressed in the time-varying canonical format.

A. SUM Operation

For inputs A and B in the time-varying canonical form, we want to represent $C = SUM(A, B)$ also in the same format. Since SUM is a linear operator, the mean and principal and random component sensitivities at the output can be obtained by adding the corresponding input sensitivities. The intrinsic component sensitivity is given by, $c_{R,int}t^{\frac{n}{2}} = \left(\sqrt{a^2_{R,int} + b^2_{R,int}}\right)t^{\frac{n}{2}}$.

B. MAX Operation

Given inputs A and B in time-varying canonical form, the aim is to find $C = MAX(A, B)$ also in the time-varying canonical format. In [25], Clark proposed an analytical model for the mean and variance (μ_c, σ_c^2) of maximum of two normally distributed random variables. In this subsection, we extend it to include the time-varying means and sensitivities.

For a timing quantity A expressed in the time-varying canonical model, the time variation of mean and sensitivities

are as in (11). Using these, we obtain the time-varying variance as,

$$\begin{aligned}\sigma_a^2(t) &= \sum_{i=1}^{N}(S_{a_{X,i0}} + a_{X,i}t^n)^2 + \sum_{j=1}^{M}(S_{a_{R,j0}} + a_{R,j}t^n)^2 \\ &\quad + a_{R,int}^2 t^n \\ &= \sigma_{a0}^2 + \Delta\sigma_a^2(t) \end{aligned} \tag{12}$$

Here, σ_{a0}^2 represents the variance of A at $t = 0$. $\Delta\sigma_a^2(t)$ denotes the deviation in variance of A from the initial value at $t = 0$ and consists of first order and second order terms. From (9), (10) and (11), the ratio of magnitude of second order to first order term is of the order of $10^{-3}t^n$ which is negligible over the lifetime of the device. Therefore, we approximate $\Delta\sigma_a^2(t)$ as,

$$\Delta\sigma_a^2(t) = \sigma_{a_t}^2 t^n \text{ where,}$$
$$\sigma_{a_t}^2 = \sum_{i=1}^{N} 2S_{a_{X,i0}}a_{X,i} + \sum_{j=1}^{M} 2S_{a_{R,j0}}a_{R,j} + a_{R,int}^2 \tag{13}$$

The tightness probability between two timing quantities A and B is defined as $\Phi(\beta)$ which is the CDF of standard normal distribution [25]. With aging, β is a function of time given by,

$$\beta(t) = \frac{\mu_a(t) - \mu_b(t)}{\theta(t)} \tag{14}$$

$$\text{where } \theta(t) = \sqrt{\sigma_a^2(t) + \sigma_b^2(t) - 2\rho(t)\sigma_a(t)\sigma_b(t)} \tag{15}$$

The correlation coefficient $\rho(t)$ captures the topological and spatial correlations between A and B at time instant t. We ignore the marginal changes in correlations between the gates due to differential stress and assume ρ to be independent of time. However, the mean and variance of *A* and *B* are functions of time thereby making β and $\Phi(\beta)$ also time-varying.

The mean and sensitivities at the output of MAX operation depend on tightness probability, which is strongly non-linear as it approaches 0 or 1. Therefore, we divide the tightness probability into four regions as shown in Fig. 1. Within each region, we can find a time-varying canonical representation for the output. In region (1), $\Phi(\beta) \approx 0$ and hence, $MAX(A, B)$ can be approximated as B. Similarly, in region (4), $\Phi(\beta) \approx 1$ and $MAX(A, B)$ approximately equals A. In these two cases, the result of the MAX operation is in the canonical form.

If β lies within region (2) or (3), the output canonical is evaluated using a first order Taylor series expansion. This is a good approximation since the variation in mean and standard deviation over the lifetime is typically small, less than 15% and 5% respectively. Assume β varies from β_i to β_j within a region in a time interval $[t_i, t_j]$. The evaluation of time-varying mean and sensitivities of the output canonical is discussed in the subsequent subsections.

1) Mean as a function of time ($\mu_c(t) = \mu_c(t_i) + \Delta\mu_c(t)$): By Clark's approximation [25], the mean delay at the output of MAX operation at $t = t_i$ is given by,

$$\mu_c(t_i) = \mu_a(t_i)\Phi(\beta_i) + \mu_b(t_i)\Phi(-\beta_i) + \theta_i\phi(\beta_i) \tag{16}$$

where $\theta_i = \theta(t_i)$. The deviation in mean from the initial value at t_i is $\Delta\mu_c(t) = c_t(t^n - t_i^n)$. Assuming small deviations, c_t is computed using first order Taylor series expansion of $\mu_c(t_i)$ with respect to the means and variances and is given by,

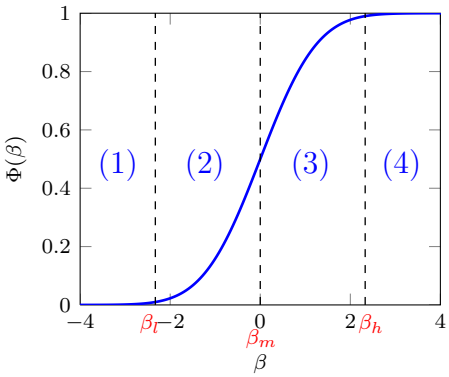


Fig. 1: The four regions in tightness probability domain determined by β_l, β_m and β_h

$$\begin{aligned} c_t &= \Phi(\beta_i)a_t + \Phi(-\beta_i)b_t \\ &\quad + \frac{\phi(\beta_i)}{2\theta_i}\Big((1 - \rho\frac{\sigma_{bi}}{\sigma_{ai}})\sigma_{a_t}^2 + (1 - \rho\frac{\sigma_{ai}}{\sigma_{bi}})\sigma_{b_t}^2\Big) \end{aligned} \tag{17}$$

where σ_{ai} and σ_{bi} denote the standard deviations of A and B at $t = t_i$ respectively.

2) Principal component sensitivity as function of time ($S_{c_{X,i}}(t) = S_{c_{X,i}}(t_i) + \Delta S_{c_{X,i}}(t)$): The sensitivity to principal component at $t = t_i$ is given by [26],

$$S_{c_{X,i}}(t_i) = S_{a_{X,i}}(t_i)\Phi(\beta_i) + S_{b_{X,i}}(t_i)\Phi(-\beta_i) \tag{18}$$

As before, the deviation in principal component sensitivity from the initial value at $t = t_i$ ($\Delta S_{c_{X,i}}(t) = c_{X,i}(t^n - t_i^n)$) is computed using first order Taylor series expansion. The variation of $\beta(t)$ within the time interval is approximated as,

$$\beta(t) \approx \beta_i + \beta_t(t^n - t_i^n) \tag{19}$$

$$\begin{aligned} \text{where, } \beta_t &= \frac{a_t - b_t}{\theta_i} + \\ &\quad \frac{(\mu_b(t_i) - \mu_a(t_i))}{2\theta_i^3}\Big((1 - \rho\frac{\sigma_{bi}}{\sigma_{ai}})\sigma_{a_t}^2 + (1 - \rho\frac{\sigma_{ai}}{\sigma_{bi}})\sigma_{b_t}^2\Big) \end{aligned} \tag{20}$$

As a result, $c_{X,i}$ is given by,

$$c_{X,i} = \Phi(\beta_i)a_{X,i} + \Phi(-\beta_i)b_{X,i} + (S_{a_{X,i}}(t_i) - S_{b_{X,i}}(t_i))\phi(\beta_i)\beta_t \tag{21}$$

The sensitivity coefficient associated with random component is derived similarly.

3) Sensitivity of intrinsic component $c_{R,int}t^{\frac{n}{2}}$: Given the sensitivities of A and B, the sensitivity associated with the intrinsic component in C is calculated as follows.

$$c_{R,int}t^{\frac{n}{2}} = t^{\frac{n}{2}}\sqrt{[a_{R,int}\Phi(\beta(t))]^2 + [b_{R,int}\Phi(-\beta(t))]^2} \tag{22}$$

Since we wish to express the output of MAX operation in time-varying canonical form, we approximate $c_{R,int}$ as,

$$c_{R,int} \approx \sqrt{(a_{R,int}\Phi(\beta_i))^2 + (b_{R,int}\Phi(-\beta_i))^2} \tag{23}$$

Therefore, each time-varying quantity is represented in the form $x(t) = x(t_i) + x_t(t^n - t_i^n)$ which is in accordance with the

canonical model defined in (11). If β increases/decreases over time and crosses over to the next region, then there is a switch from one canonical model to the other and the corresponding break point in time is evaluated using (19). Thus, the output of MAX operator is stored as a collection of time-varying canonicals with different canonicals being valid over different time intervals.

C. Implementation

Let the inputs to MAX operation be A_c and B_c where A_c and B_c are a collection of time-varying canonicals with arbitrary number of break points and defined over the lifetime (t_{life}). Let $t_{A1}, t_{A2}, \cdots t_{life}$ be the break points in A_c and $t_{B1}, t_{B2}, \cdots t_{life}$ be the break points in B_c. In the first step, the break points in A_c and B_c together are sorted in the increasing order to find the break points in the output of MAX operation. The input canonicals associated with each of these time intervals are also identified. For a given time interval $[t_i, t_j]$ with input canonicals A_i and B_i, the identification of break points within this interval and the evaluation of MAX is outlined in Algorithm 1.

We start with defining β_l, β_m and β_h, the points which determine the four regions in $\Phi(\beta)$ curve. β_m is fixed at zero and β_l and β_h correspond to the points where $\Phi(\beta)$ is 0.01 and 0.99 respectively [Step 1]. The next step is to identify if there is any switching of regions during the time interval $[t_i, t_j]$ over which the inputs are defined. The values of β_l, β_m and β_h are compared against β_i and β_j to check if there is any switching of regions. If so, the time at which switching of regions occur is evaluated and this is added to the list of break points. The list of break points is then sorted in ascending order and time t_j is added as the last break point [Steps 2-13].

Once the break points are obtained, we evaluate the output canonical in each of the identified time intervals [Steps 14-30]. This is done as discussed in the previous section. If $\beta \leq \beta_l$ in the time interval under consideration, the output canonical is the same as input B_i. Similarly, if $\beta \geq \beta_h$ in the time interval, the output canonical equals A_i. If $\beta_l < \beta \leq \beta_m$ or $\beta_m < \beta < \beta_h$ in the time interval, the output canonical is computed using the formulas derived in sections (III-B1) to (III-B3).

To keep a limit on the number of break points, a check is performed after every MAX operation to see if any of the neighbouring canonicals can be merged without introducing significant error. Let X and Y be two neighbouring canonicals with time intervals $[t_1, t_2]$ and $[t_2, t_3]$ respectively. The two canonicals are merged to obtain canonical X valid over time interval $[t_1, t_3]$ if the error in mean and standard deviation at t_3 introduced in this approximation is within 1%.

Algorithm 1: Maximum of two time-varying canonicals

Input : A_i and B_i
Output: $C_i = MAX(A_i, B_i)$
1: $\beta_l \leftarrow -2.33$, $\beta_m \leftarrow 0$ and $\beta_h \leftarrow 2.33$
2: $\beta_i \leftarrow \beta(t = t_i)$
3: Find β_t using (20)
4: $\beta_j \leftarrow \beta_i + \beta_t(t_j^n - t_i^n)$
5: Initialise $\beta_{brk} = [\beta_l, \beta_m, \beta_h]$; t_{brk} = [ ]
6: **for** $\beta_b \in \beta_{brk}$ **do**
7: **if** $\beta_i < \beta_b < \beta_j$ **then**
8: Solve $\beta_b = \beta_i + \beta_t(t_b^n - t_i^n)$ to find t_b
9: Add t_b to t_{brk}
10: **end if**
11: **end for**
12: Sort t_{brk} in ascending order
13: Add t_j to t_{brk}
14: $k \leftarrow 0$
15: Initialise C_i =[ ]
16: **while** $k < len(t_{brk})$ **do**
17: **if** $k == 0$ **then**
18: Time interval, TI = [t_i,$t_{brk}[k]$]
19: **else**
20: Time interval, TI = $[t_{brk}[k-1], t_{brk}[k]]$
21: **end if**
22: **if** $\beta \leq \beta_l$ in TI **then**
23: Add B_i to C_i
24: **else if** $\beta \geq \beta_h$ in TI **then**
25: Add A_i to C_i
26: **else**
27: Evaluate output canonical using formulas derived in Sections (III-B1) to (III-B3) and add to C_i
28: **end if**
29: $k+ = 1$
30: **end while**
31: return C_i

IV. Time-Varying Path Criticality

The global criticality (Q_p) of path p is defined as the probability that delay of the path (D_p) exceeds the complementary path delay (CD_p) [15]. With aging, all the gate delays vary with time, thereby making the path delay and complementary path delay also time-varying. Thus the global criticality of a path under aging is a function of time that is given by,

$$Q_p(t) = P(D_p(t) > CD_p(t)) \tag{24}$$

where $D_p(t)$ and $CD_p(t)$ are Gaussian random variables with time-varying means and variances. A path p in the circuit is treated as a potential critical path if its maximum global criticality over the lifetime exceeds a threshold, *i.e.*

$$\max_{0 \leq t \leq t_{life}} \left\{ Q_p(t) \right\} \geq \gamma \tag{25}$$

where t_{life} is the expected lifetime and γ is the criticality threshold. This requires the evaluation of maximum global criticality of a path over the lifetime of the circuit. However, the global path criticality can vary non-monotonically over time as indicated in Fig. 2. These criticality values are obtained using Monte Carlo simulations from 0 to 10 years in steps of one year.

It is infeasible to compute the criticality at each time instant within the lifetime for determining whether the path is potentially critical or not. We derive the following properties of time-varying path criticalities under aging to overcome this problem.

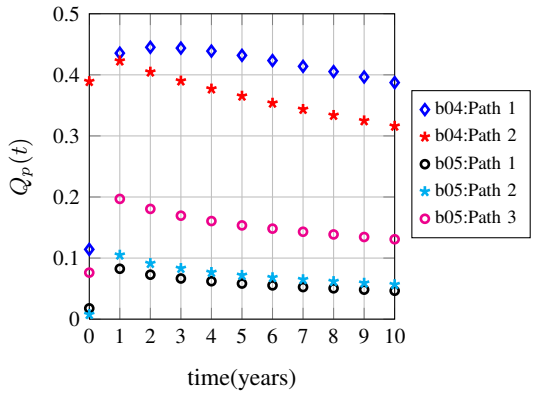


Fig. 2: Non-monotonic time-varying path criticalities of some paths in b04 and b05 circuits obtained using Monte Carlo simulations in 90 nm technology.

Property 1. *If $D_p(t) = \mathcal{N}(\mu_{D_p}(t), \sigma^2_{D_p}(t))$ and $CD_p(t) = \mathcal{N}(\mu_{CD_p}(t), \sigma^2_{CD_p}(t))$ denote the time-varying path delay and complementary path delay respectively, then the time-varying global path criticality is given by $\Phi(\beta(t))$ where*

$$\beta(t) = \frac{\mu_{D_p}(t) - \mu_{CD_p}(t)}{\sqrt{\sigma^2_{D_p}(t) + \sigma^2_{CD_p}(t) - 2\rho(t)\sigma_{D_p}(t)\sigma_{CD_p}(t)}} \tag{26}$$

and $\rho(t)$ is the correlation coefficient between path delay and complementary delay at time t.

Proof. The time-varying global path criticality is given by,

$$Q_p(t) = P(D_p(t) > CD_p(t)) \tag{27}$$

Based on the delay model given by (11), $D_p(t)$ and $CD_p(t)$ are jointly Gaussian random variables with time-varying means, variances and correlation coefficient. Let $D(t)$ denote the delay difference random variable $D_p(t) - CD_p(t)$. Therefore, at any given time t,

$$Q_p(t) = P(D(t) > 0) = \Phi\Big(\frac{\mu_d(t)}{\sigma_d(t)}\Big) = \Phi(\beta(t)) \tag{28}$$

where,

$$\mu_d(t) = \mu_{D_p}(t) - \mu_{CD_p}(t) \text{ and} \tag{29}$$

$$\sigma_d(t) = \sqrt{\sigma^2_{D_p}(t) + \sigma^2_{CD_p}(t) - 2\rho(t)\sigma_{D_p}(t)\sigma_{CD_p}(t)} \tag{30}$$

□

Property 2. *If $A_i(t)$ and $B_i(t)$ are two time-varying canonicals valid over the time interval $[t_i, t_j]$ with β as in (19), then criticality of A_i with respect to B_i will be maximum either at t_i or t_j.*

Proof. Let $Q_{A_iB_i}(t)$ denote the criticality of A_i with respect to B_i at any time instant $t \in [t_i, t_j]$. We have to show that,

$$\max_{t_i \le t \le t_j} \{Q_{A_iB_i}(t)\} = \max\{Q_{A_iB_i}(t_i), Q_{A_iB_i}(t_j)\} \tag{31}$$

Given that $A_i(t)$ and $B_i(t)$ are represented in the time-varying canonical form with $\beta(t)$ as in (19).

$$\text{If } \beta_t \ge 0, \max_{t_i \le t \le t_j} \{\beta(t)\} = \beta(t_j)$$

$$\text{If } \beta_t < 0, \max_{t_i \le t \le t_j} \{\beta(t)\} = \beta(t_i)$$

$$\begin{aligned}\max_{t_i \le t \le t_j} \{Q_{A_iB_i}(t)\} &= \max_{t_i \le t \le t_j} \{\Phi(\beta(t))\} \\ &= \max\{\Phi(\beta(t_i)), \Phi(\beta(t_j))\} \\ &= \max\{Q_{A_iB_i}(t_i), Q_{A_iB_i}(t_j)\}\end{aligned}$$

□

Property 3. *If $A(t)$ and $B(t)$ are a collection of time-varying canonicals over the lifetime, the break points in $\max\{A(t), B(t)\}$ will be a super set of the union of break points in $A(t)$ and $B(t)$.*

Proof. Let $t_b(A)$ be the break points in $A(t)$ and $t_b(B)$ be the break points in $B(t)$. Let $t_{A\cup B} = t_b(A) \cup t_b(B)$ denote the union of break points in $A(t)$ and $B(t)$. If $C(t) = \max(A(t), B(t))$, we want to show that $t_b(C) \supseteq t_{A\cup B}$.

According to the implementation details in Section (III-C), MAX operation only introduces additional break points. Therefore, break points in $C(t)$ will always be a superset of the break points in $A(t)$ and $B(t)$, *i.e.* $t_b(C) \supseteq t_{A\cup B}$. □

Property 4. *If $A(t)$ and $B(t)$ are a collection of time-varying canonicals over the lifetime, then the criticality of A with respect to B will be maximum either at $t = 0$ or at the break points of $\max\{A(t), B(t)\}$.*

Proof. Let $Q_{AB}(t)$ denote the criticality of A with respect to B at any time instant $t \in [0, t_{life}]$. We have to show that

$$\max_{0 \le t \le t_{life}} \{Q_{AB}(t)\} = \max_{t_i \in t_{bmax}} \{Q_{AB}(0), Q_{AB}(t_i)\} \tag{32}$$

where t_{bmax} denote the break points in $\max\{A(t), B(t)\}$. This property follows directly from properties 2 and 3. □

Although this limits the points at which the criticality needs to be computed, it can be still very large as the set of break points for each path delay and its corresponding complementary path delay could be different. The following property reduces this set significantly.

Property 5. *The time-varying global criticality of any path will be maximum either at t = 0 or at any of the break points in circuit delay distribution.*

Proof. Consider a path i in the circuit and let $D_{p,i}(t)$ and $CD_{p,i}(t)$ denote its delay and complementary path delay respectively. If there are N paths in the circuit, complementary path delay is given by,

$$CD_{p,i}(t) = \max_{\substack{1 \le j \le N \\ j \ne i}} \{D_{p,j}(t)\} \tag{33}$$

The circuit delay distribution (D_{ckt}) is the statistical maximum of the delay of all the paths in the circuit. Therefore,

$$D_{ckt}(t) = \max\{D_{p,i}(t), CD_{p,i}(t)\} \tag{34}$$

Making use of (34) and property 4, we get

$$\max_{0\le t\le t_{life}}\{Q_{p,i}(t)\} = \max_{t_i\in t_b(D_{ckt})}\{Q_{p,i}(0), Q_{p,i}(t_i)\} \quad (35)$$

where $t_b(D_{ckt})$ denotes the break points in the circuit delay distribution. This holds true for any path i in the circuit. Therefore, the global criticality of any path in the circuit will be maximum either at t = 0 or at the break points in the circuit delay distribution. □

As a result of this property, we need to now evaluate criticalities and identify PCPs only at $t = 0$ and at the break points of circuit delay distribution. This typically is less than four as will be seen in the results section.

V. Statistical PCP Selection

The objective of the work is to identify the set of potential critical paths (S_{PCP}) with maximum global criticality over the lifetime greater than a threshold (γ). Mathematically,

$$\text{A path } p \in S_{PCP} \text{ if } \max_{0\le t\le t_{life}}\left\{Q_p(t)\right\} \ge \gamma \quad (36)$$

By property 5 given in the previous section, the global path criticality will be maximum either at t = 0 or at the break points in the circuit delay distribution. Therefore, we can conclude that a search for critical paths at these time points is sufficient to identify the potential critical paths under aging.

We make use of the hierarchical partitioning algorithm proposed in [15] to identify the paths with global criticality above a specified threshold at a given time instant. This algorithm integrates path selection and global criticality computation. This is done hierarchically with the circuit being split into disjoint groups at each level. The extension of hierarchical partitioning algorithm for identification of potential critical paths under NBTI is outlined in Algorithm 2.

The inputs to the potential critical path selection algorithm are the circuit graph (G) and the desired criticality threshold (γ). In order to take into account the errors in the time-varying SSTA formulation and path criticality computation, we choose the threshold for hierarchical partitioning to be lower than the desired threshold depending on the average error value (Δ). We also go through two iterations. In the first iteration, we identify the paths with criticality above the threshold at t = 0 and at the break points in circuit delay distribution using the hierarchical partitioning algorithm. Once this is done, we compare the sum of path criticalities at each of the break points against the value at t = 0. If the value has dropped significantly, it is possible that we have missed some critical paths at that break point. This could happen due to the errors inherent in the SSTA formulation. Therefore, we reduce the threshold further and go through a second iteration to identify an additional set of critical paths.

A detailed description of the algorithm is given below. We start with a time-varying forward SSTA to obtain the circuit delay distribution as a function of time. The break points in the circuit delay along with $t = 0$ form the time instants (T_{brk}) at which hierarchical partitioning should be done for the PCP selection [Steps 1-2]. In the next step, we perform the first level of partitioning based on output nets [Steps 3-5]. All the paths passing through a given output net i form a level (1) group and the arrival time at the output net ($AT_i(t)$) indicates its time-varying delay distribution.

A level (1) group G_k needs to be partitioned further at any time instant t_i only if its global criticality $Q_{g,k}(t_i)$ exceeds the threshold γ. This is because, the criticality of any path within the group cannot be greater than the group criticality. For each time instant t_i, a pruning based on local criticality is performed to remove the non-dominant groups whose criticality is guaranteed to be lower than the threshold [Step 8]. A clustering of similar groups is also done to merge groups that have high topological correlation as discussed in [15]. The criticality of the dominant level (1) groups are then evaluated [Steps 12-13]. If the criticality of group G_k exceeds the threshold at t_i, then group G_k should be partitioned at t_i. This is done in two iterations as explained earlier. In the first iteration [Steps 15-21], the threshold is set at a value Δ below the desired value and the set of critical paths at t_i are

Algorithm 2: PCP selection using hierarchical partitioning

Input : $G(V, E)$: Graph representing circuit netlist
γ : Criticality threshold
Output: S_{PCP} : Set of potential critical paths

/* Time-varying SSTA */
1: Perform time-varying SSTA to get circuit delay $D_{ckt}(t)$
2: $T_{brk} \leftarrow \{0\} \cup t_b(D_{ckt})$
/* Level (1) partitioning based on output nets */
3: **for** each net $i \in$ output nets N_{out} **do**
4: $D_i(t) \leftarrow AT_i(t)$
5: **end for**
/* Hierarchical partitioning of critical level (1) groups */
6: **for** each time $t_i \in T_{brk}$ **do**
7: Find gate delays and arrival times by substituting $t = t_i$ in the time-varying canonical form
8: $G_{dom}(t_i) \leftarrow$ Dominant groups at t_i after pruning and clustering of similar groups
9: $Qlist \leftarrow \{\}$ //Initialize list of group criticalities at t_i
10: $\Sigma Q_P(t_i) \leftarrow 0$ //Initialize sum of path criticalities at t_i
11: **for** each group $G_k \in G_{dom}(t_i)$ **do**
12: Evaluate complementary delay $CD_{g,k}(t_i)$
13: $Q_{g,k}(t_i) = P(D_{g,k}(t_i) > CD_{g,k}(t_i))$
14: Add $Q_{g,k}(t_i)$ to $Qlist$
/* Iteration : 1 */
15: $\gamma_1 = \gamma - \Delta$
16: **if** $Q_{g,k}(t_i) \ge \gamma_1$ **then**
17: $S_P(t_i) \leftarrow$ Hier_partitioning(G_k,t_i,$Q_{g,k}(t_i)$,γ_1)
18: $S_{PCP} \leftarrow S_{PCP} \cup S_P(t_i)$
19: $Q_P(t_i) \leftarrow$ Sum of criticalities of paths in $S_P(t_i)$
20: $\Sigma Q_P(t_i) += Q_P(t_i)$
21: **end if**
22: **end for**
/* Iteration : 2 */
23: **if** $t_i \neq 0$ and $\Sigma Q_P(0) - \Sigma Q_P(t_i) > \Delta$ **then**
24: Lower the threshold to $\gamma_2 = \gamma - 2\Delta$
25: **for** each group $G_k \in G_{dom}(t_i)$ **do**
26: $Q_{g,k}(t_i) \leftarrow$ Criticality of G_k at t_i from $Qlist$
27: **if** $Q_{g,k}(t_i) \ge \gamma_2$ **then**
28: $S_P(t_i) \leftarrow$ Hier_partitioning(G_k,t_i,$Q_{g,k}(t_i)$,γ_2)
29: $S_{PCP} \leftarrow S_{PCP} \cup S_P(t_i)$
30: **end if**
31: **end for**
32: **end if**
33: **end for**

obtained. The threshold Δ is chosen based on the average error in path criticality computation. If the sum of path criticalities has dropped by Δ in comparison to the value at $t = 0$, a second round of hierarchical partitioning is done after lowering the threshold further [Steps 23-32]. Finally, the set of potential critical paths is the union of critical paths identified at each of these partitioning steps.

VI. Proposed Framework

The various steps involved in the proposed framework is given in Fig. 3. The first step is to evaluate the threshold voltage degradation parameters A_0 and S_A as defined in Section II(A). These are constants for a given technology and operating conditions. Following this, circuit synthesis is performed to obtained the synthesized netlist. A place and route is also performed to obtain the location information of the gates. Given the synthesized netlist, a logic simulation is performed using 10^5 input vectors to generate the Value Change Dump (VCD) file. The VCD file contains information regarding the signal transitions at the various nodes. This is processed to obtain the static signal probabilities and the correlation coefficients which are then used to evaluate the stress factors for the various PMOS transistors.

Once the stress factors has been computed, each of the edge delays is represented using the time-varying canonical delay model defined in (11). A time-varying forward SSTA is then performed to obtain the circuit delay distribution as a collection of time-varying canonicals with break points in time. The MAX operation on time-varying canonicals is done as described in Algorithm 1. An extension of the hierarchical partitioning algorithm is then employed at $t = 0$ and at the break points of the circuit delay distribution to identify the PCPs under aging. This is done as described in Algorithm 2.

VII. Results

The proposed algorithm for time-varying SSTA and potential critical path selection is implemented in C++ and the experiments are performed on ISCAS'85, ISCAS'89 and ITC'99 benchmark circuits synthesized using UMC 90nm and NanGate 45nm [27] technology libraries. The variations in 3 process parameters, L, W and V_{th}, are considered along with NBTI for timing analysis. The standard deviation of each parameter variation is assumed to be 10% of the nominal value and 5% random variation is assumed. A three level quad-tree is used for modelling the spatial correlations [28] and the variance is divided equally among the three layers. We aim to identify the potential critical paths with maximum global criticality above 0.05. From the experiments performed on various circuits to identify the critical paths at $t = 0$ using the hierarchical partitioning technique in [15], it was found that the average error in criticality is around 0.02 for both UMC 90nm and NanGate 45nm technologies. Therefore, the threshold Δ is fixed at 0.02. Accordingly, the criticality threshold is first set to 0.03 and is lowered to 0.01 for the second iteration. The results are verified by comparing against Monte Carlo simulations considering 10^5 samples. The circuit lifetime is assumed to be 10 years.

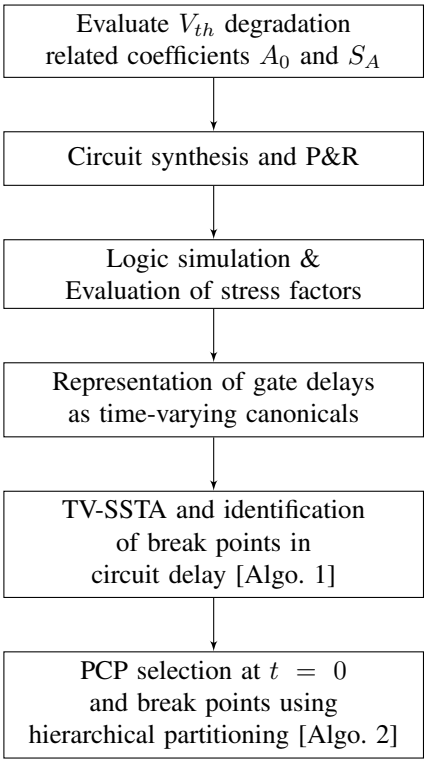
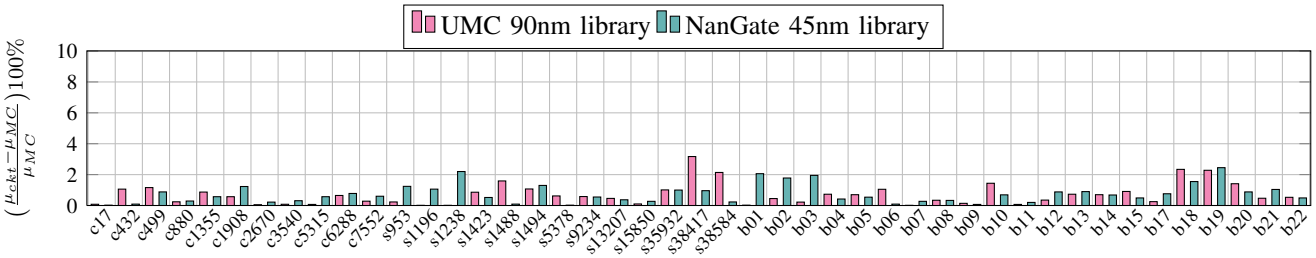


Fig. 3: Steps in the proposed framework

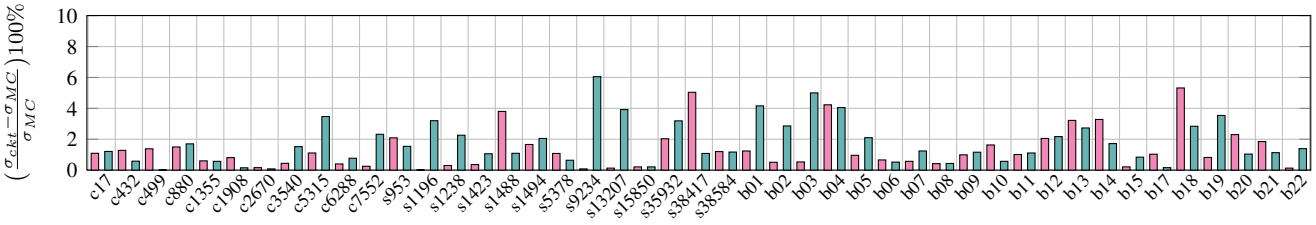
A. Time-varying SSTA

Fig. 4 shows the percentage error in mean and standard deviation of circuit delay at the end of lifetime using the proposed time-varying SSTA. The maximum error in mean of 3.17% occurs for s38417 circuit in UMC 90nm technology and the average error is 0.70%. The maximum error in standard deviation is around 6% occurring in s9234 circuit synthesized using NanGate 45nm library and the average error is 1.48%. If a conventional SSTA is performed using aged gate delays at $t = 10$ years, the error in mean is found to be 5.06% in s38417 circuit. Similarly, the error in standard deviation for s9234 circuit is 5.00%. This indicates that the error in these circuits is primarily due to linearization of MAX operation in the standard SSTA formulation and there is no significant additional error due to the time-varying formulation.

Fig. 5 compares the time variation in mean and standard deviation of circuit delay against the values obtained by re-running conventional SSTA multiple times, from 0 to 10 years in steps of 1 year for three circuits, c432, s13207 and b04, in UMC 90nm technology. The mean of circuit delay is plotted as a function of time in Fig. 5(a). The figure indicates that the mean of circuit delay increases with time and matches closely with the values obtained from individual SSTA runs. The error in mean of circuit delay at 10 years is found to be 0.02%, 0.07% and 0.13% for c432, s13207 and b04 circuits respectively. The time variation in standard deviation is plotted in Fig. 5(b). There is a discontinuity of 1.5% in standard deviation of s13207 which occurs at a break point in the circuit delay distribution. This is seen in some circuits and is typically less than 1%. The error in standard deviation of circuit delay

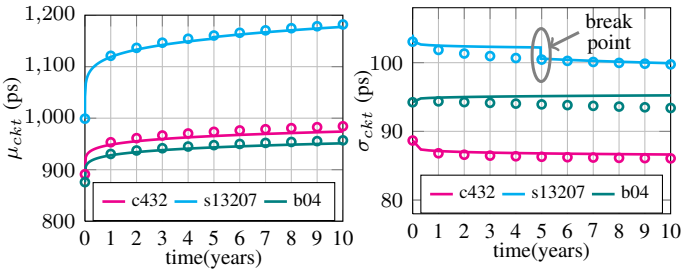


(a) Percentage error in mean of circuit delay



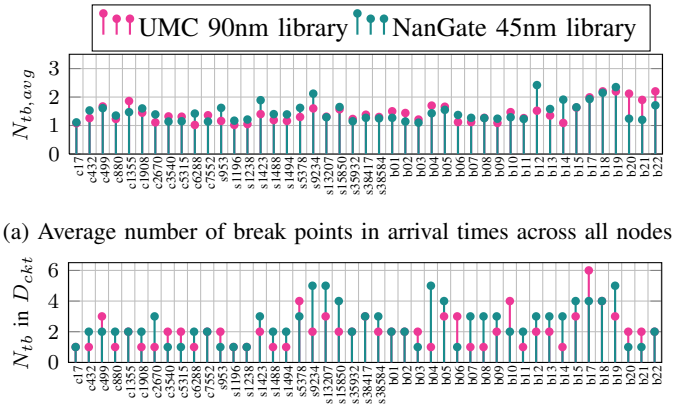
(b) Percentage error in standard deviation of circuit delay

Fig. 4: Percentage error in mean and standard deviation of circuit delay at $t = 10$ years with respect to Monte Carlo simulations using the proposed time-varying SSTA technique



(a) Mean (b) Standard Deviation

Fig. 5: Time variation in mean and standard deviation of circuit delay obtained using the proposed method in UMC 90nm technology. The 'o' markers denote the values obtained using conventional SSTA runs.



(a) Average number of break points in arrival times across all nodes

(b) Number of break points in circuit delay

Fig. 6: Number of break points using time-varying SSTA

at 10 years is found to be 0.62%, 0.04% and 2.22% for c432, s13207 and b04 circuits respectively.

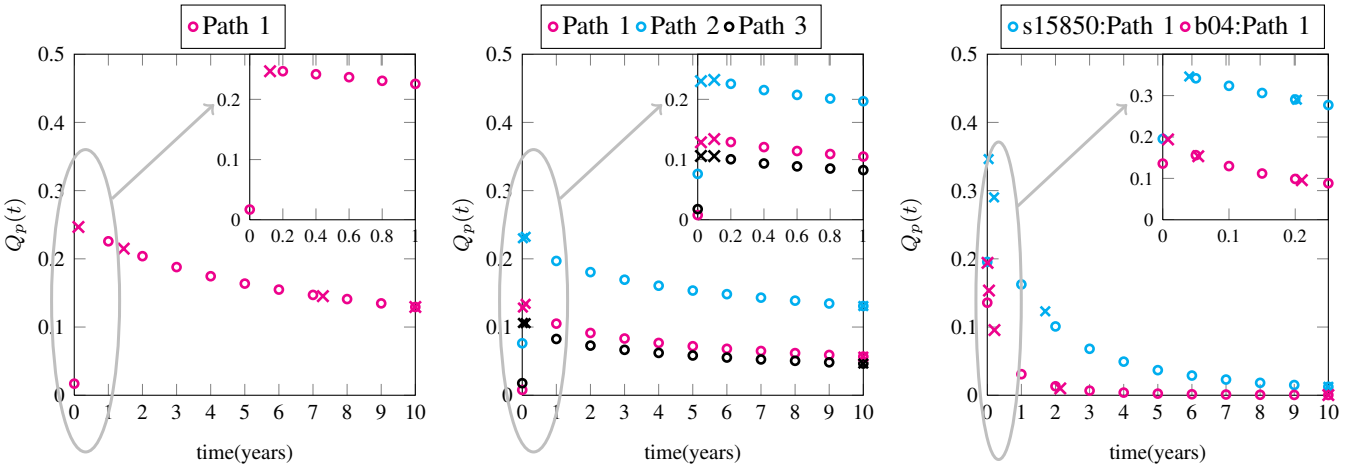
The average number of break points in arrival times across all the nodes ($N_{tb,avg}$) is shown in Fig. 6(a) and is found to be less than 2 in most of the benchmark circuits. This determines the run time complexity of the time-varying SSTA algorithm and equals $O(N_{tb,avg}(|V|+|E|))$ which is $N_{tb,avg}$ times that of a conventional SSTA. On an average, the proposed SSTA algorithm is able to predict the time-varying circuit delay distribution in about 2x times the time required for a single SSTA run, with a maximum of around 3x for larger circuits.

The number of break points (N_{tb}) in circuit delay distribution is given in Fig. 6(b). The maximum number of break points is 6 in the case of b17 circuit synthesized using UMC 90nm technology and is less than 4 for most of the benchmark circuits. These break points determine the time instants at which criticality evaluation and PCP selection will be done.

B. Time-varying Path Criticality

To validate Property 5 in Section IV, we have considered a few circuits where there exists at least one path with non-monotonic criticality. Monte Carlo simulations are performed at the break points in the circuit delay distribution and some test points and the criticality values are compared as shown in Fig. 7. The 'x' markers in the figure denote the path criticalities obtained using Monte Carlo simulations at the break points. The 'o' markers denote the path criticalities obtained using Monte Carlo simulations at the test points $t \in \{0, 1, \cdots, 10\}$ years.

In s5378 circuit synthesized using UMC 90nm technology, there are four break points in the circuit delay distribution at t = 0.12, 1.45, 7.29 and 10 years, as seen in Fig. 7(a). From the figure, we can see that the maximum criticality for the path occurs at some time instant between 0 and 1 year.



(a) s5378 circuit (b) b05 circuit (c) s15850 and b04 circuits

Fig. 7: Time-varying path criticality in s5378 (90 nm), b05 (90 nm) and s15850 and b04 (45 nm) circuits. The 'x' and 'o' markers denote the Monte Carlo path criticalities at the break points in the circuit delay and the test points respectively. A zoomed version of the region where maximum path criticality occurs is given in the inset figure.

Therefore, we have considered additional test points in this region, $t \in \{0.2, 0.4 \cdots, 0.8\}$ years and the results are shown in the inset figure. From the results, we can conclude that the criticality of path 1 is maximum at the break point, t = 0.12 years. Similarly, in b05 circuit, there are three break points in the circuit delay distribution at t = 0.02, 0.1 and 10 years. The inset figure indicates that the maximum criticality for all the paths occur at the break point t = 0.1 years. Similar results have also been obtained for circuits synthesized using NanGate 45nm library as seen in the case of s15850 and b04 circuits in Fig. 7(c).

It should be noted that the path criticality values plotted in Fig. 7 are obtained using Monte Carlo simulations. If we use time-varying SSTA along with hierarchical partitioning to find the path criticalities, it is true that with existing SSTA based methods, it will be difficult to capture the time variation exactly. But, as shown in Fig. 8, the overall trend is clear in most cases.

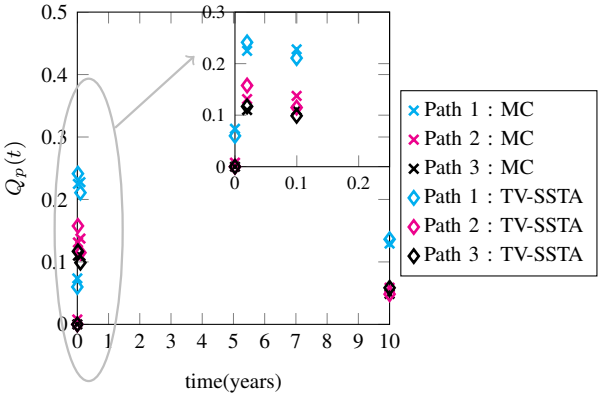
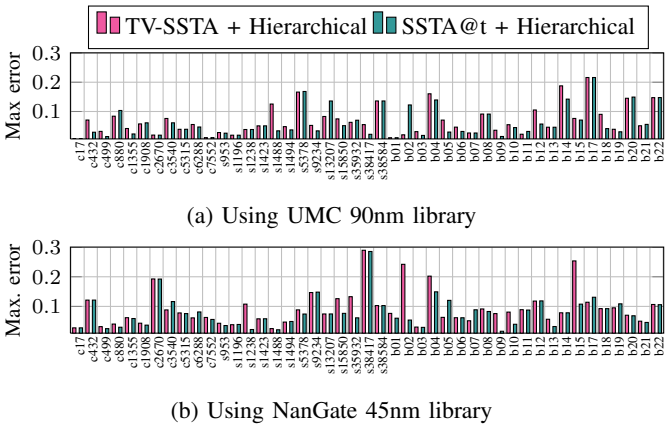


Fig. 8: Time-varying path criticality in b05 circuit synthesized using UMC 90nm library. The 'x' and '◇' markers denote the path criticalities at the break points in circuit delay computed using Monte Carlo and hierarchical methods respectively.



(a) Using UMC 90nm library

(b) Using NanGate 45nm library

Fig. 9: Maximum error in path criticality with respect to Monte Carlo simulations for (a) Time-varying SSTA + Hierarchical partitioning and (b) SSTA at a given time point + Hierarchical partitioning in various benchmark circuits

C. Statistical PCP Selection

Fig. 9 contains a comparison of the maximum error in path criticality computed using the proposed method and by performing a conventional SSTA followed by criticality computation at the break points. The aim is to see if any additional error is introduced due to the approximations made in the time-varying SSTA formulation. The results show that the maximum error in both the cases is comparable, both for UMC 90nm and NanGate 45nm technologies, indicating that the approximations involved in the time-varying SSTA formulation does not introduce any significant error in the path criticality. The maximum error in path criticality is between 0.2 and 0.3 and is found to be less than 0.1 in most of the benchmark circuits. Also, the average error in path criticality across all the break points equals 0.02 in both UMC 90nm and NanGate 45nm technologies.

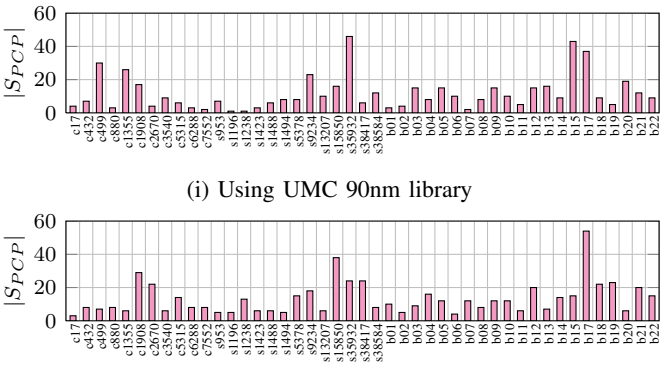
TABLE I: Path criticality in circuits where the maximum error is above 0.1 using UMC 90nm technology

Ckt	Max error	Time	Q_{hier}	Q_{MC}
s1488	0.1255	10 yr	0.2552	0.3807
s5378	0.1663	10 yr	0.3899	0.5563
s38584	0.1360	0	0.2650	0.4010
b04	0.1605	10 yr	0.1784	0.0178
b14	0.1876	10 yr	0.0272	0.2148
b17	0.2161	0	0.2243	0.4403
b20	0.1452	4.11 yr	0.2934	0.4386
b22	0.1474	10 yr	0.2256	0.3730

The error in path criticality computations will affect the identification of potential critical paths. It should be noted that it is not just the maximum error, even a small error can push the criticality below the threshold. This is the reason why we have a guard band. Table I contains a list of circuits synthesized using UMC 90nm technology in which the maximum error in path criticality is above 0.1. The time point at which the maximum error occurs along with the criticality values obtained using the proposed method (Q_{hier}) and Monte Carlo simulations (Q_{MC}) are also given in the table. The results indicate that in most cases the actual criticality values are well above the threshold so that the error does not cause the criticality to fall below the threshold. The only issue is in b14 circuit and in this case also, the path is identified when the threshold is lowered to 0.01.

To check the impact of error in path criticality on the identification of potential critical paths, we identified PCPs using both Monte Carlo and hierarchical methods and the comparison is shown in Fig. 10. In both the methods, the potential critical paths are identified by performing criticality computations at $t = 0$ and at the break points in the circuit delay distribution. Fig. 10(a) contains the number of PCPs obtained using the hierarchical method and Fig. 10(b) contains the number of PCPs with Monte Carlo criticality above 0.05 not identified by the proposed method in spite of lowering the threshold. It is a stacked plot indicating the maximum criticality of each of the missed paths. Note that the figure contains only those circuits in which the proposed method has failed to identify some PCPs. The results indicate that, in UMC 90nm technology, a total of 8 paths are missed by the proposed method with a maximum of 3 paths in c432 circuit. The maximum criticality missed is 0.08 in the case of b15 circuit. A similar trend has also been observed in the case of circuits synthesized using NanGate 45nm library. Here, the number of paths missed by the proposed method is 10 with a maximum of 3 paths in c5315 circuit. The maximum criticality of the path missed is 0.09 in the case of c2670 circuit.

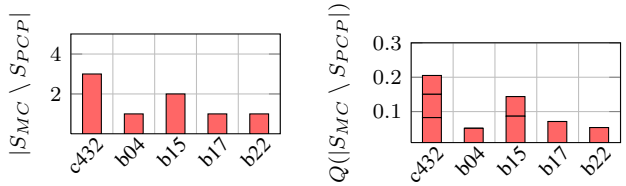
In order to check if the number of paths reported by our method is reasonable, we did a Monte Carlo analysis with three different thresholds (0.05, 0.03 and 0.01) and the results are given in Fig. 11. Note that in all the circuits, the number of paths identified using our method is greater than or equal to the number of paths reported by Monte Carlo analysis with the threshold set to 0.05. In our method, since we lower the threshold to compensate for the errors, we report a larger number of paths. This is especially seen in some circuits like c1355, s35932 and b15. This is because there are a large number of paths in these circuits with criticality in and around



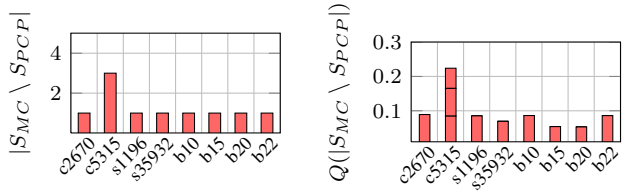
(i) Using UMC 90nm library

(ii) Using NanGate 45nm library

(a) Number of potential critical paths reported by proposed method



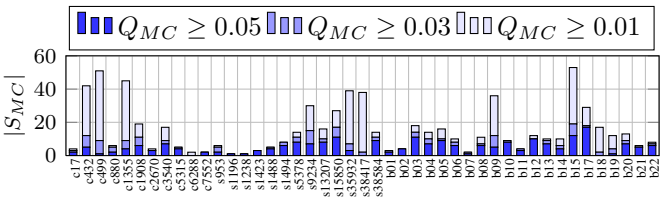
(i) Using UMC 90nm library



(ii) Using NanGate 45nm library

(b) Number of potential critical paths not identified by the proposed method with the associated maximum criticalities

Fig. 10: Comparsion of paths reported by the proposed method against Monte Carlo simulations



(a) Using UMC 90nm library

(b) Using NanGate 45nm library

Fig. 11: Number of potential critical paths identified by Monte Carlo simulations for thresholds of 0.05, 0.03 and 0.01

the threshold and we end up identifying all these paths.

VIII. Conclusion

In this paper, we have proposed a framework for time-varying SSTA wherein the circuit delay distribution is obtained as a collection of time-varying canonicals valid over different periods with break points in time. We proved that the global criticality of any path in the circuit will be maximum either at $t = 0$ or at the break points in circuit delay distribution. This has also been verified using Monte Carlo simulations. As a result, the criticality computation and PCP selection needs to be done only at these time points.

We have integrated the time-varying SSTA along with a path selection and criticality computation technique proposed earlier to identify the PCPs under aging. Simulation results indicate that the proposed technique is able to identify the potential critical paths within a range of the specified threshold.

References

[1] M. Agarwal, B. C. Paul, M. Zhang, and S. Mitra, "Circuit failure prediction and its application to transistor aging," in *25th IEEE VLSI Test Symposium (VTS'07)*, pp. 277–286, May 2007.

[2] S. Wang, M. Tehranipoor, and L. Winemberg, "In-field aging measurement and calibration for power-performance optimization," in *2011 48th ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp. 706–711, June 2011.

[3] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "Adaptive techniques for overcoming performance degradation due to aging in cmos circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, pp. 603–614, April 2011.

[4] W. Wang, Z. Wei, S. Yang, and Y. Cao, "An efficient method to identify critical gates under circuit aging," in *2007 IEEE/ACM International Conference on Computer-Aided Design*, pp. 735–740, Nov 2007.

[5] J. Chen, S. Wang, and M. Tehranipoor, "Efficient selection and analysis of critical-reliability paths and gates," in *Proceedings of the Great Lakes Symposium on VLSI*, GLSVLSI '12, (New York, NY, USA), pp. 45–50, ACM, 2012.

[6] A. H. Baba and S. Mitra, "Testing for transistor aging," in *2009 27th IEEE VLSI Test Symposium*, pp. 215–220, May 2009.

[7] J. Vazquez-Hernandez, "Error prediction and detection methodologies for reliable circuit operation under nbti," in *2014 International Test Conference*, pp. 1–10, Oct 2014.

[8] A. F. Gomez and V. Champac, "Early selection of critical paths for reliable nbti aging-delay monitoring," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 24, pp. 2438–2448, July 2016.

[9] A. F. Gomez and V. Champac, "Selection of critical paths for reliable frequency scaling under bti-aging considering workload uncertainty and process variations effects," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 23, pp. 27:1–27:21, Feb. 2018.

[10] D. Lorenz, M. Barke, and U. Schlichtmann, "Monitoring of aging in integrated circuits by identifying possible critical paths," *Microelectronics Reliability*, vol. 54, no. 6, pp. 1075 – 1082, 2014.

[11] H. D. Mogal, H. Qian, S. S. Sapatnekar, and K. Bazargan, "Fast and accurate statistical criticality computation under process variations," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems.*, vol. 28, pp. 350–363, March 2009.

[12] Y. Lu, L. Shang, H. Zhou, H. Zhu, F. Yang, and X. Zeng, "Statistical reliability analysis under process variation and aging effects," in *2009 46th ACM/IEEE Design Automation Conference*, pp. 514–519, July 2009.

[13] J. Fang and S. S. Sapatnekar, "The impact of bti variations on timing in digital logic circuits," *IEEE Transactions on Device and Materials Reliability*, vol. 13, pp. 277–286, March 2013.

[14] F. Firouzi, S. Kiamehr, and M. B. Tahoori, "Statistical analysis of bti in the presence of process-induced voltage and temperature variations," in *2013 18th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 594–600, Jan 2013.

[15] P. R. Chithira and V. Vasudevan, "A hierarchical technique for statistical path selection and criticality computation," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 23, pp. 9:1–9:24, Aug. 2017.

[16] S. Bhardwaj, W. Wang, R. Vattikonda, Y. Cao, and S. B. K. Vrudhula, "Predictive modeling of the nbti effect for reliable design.," in *CICC*, pp. 189–192, 2006.

[17] W. Wang, V. Reddy, B. Yang, V. Balakrishnan, S. Krishnan, and Y. Cao, "Statistical prediction of circuit aging under process variations," in *2008 IEEE Custom Integrated Circuits Conference*, pp. 13–16, Sept 2008.

[18] K. Kang, S. P. Park, K. Roy, and M. A. Alam, "Estimation of statistical variation in temporal nbti degradation and its impact on lifetime circuit performance," in *2007 IEEE/ACM International Conference on Computer-Aided Design*, pp. 730–734, Nov 2007.

[19] S. Han and J. Kim, "Nbti-aware statistical timing analysis framework," in *23rd IEEE International SOC Conference*, pp. 158–163, Sept 2010.

[20] T. Siddiqua, S. Gurumurthi, and M. R. Stan, "Modeling and analyzing nbti in the presence of process variation," in *2011 12th International Symposium on Quality Electronic Design*, pp. 1–8, March 2011.

[21] S. Han, J. Choung, B. S. Kim, B. H. Lee, H. Choi, and J. Kim, "Statistical aging analysis with process variation consideration," in *2011 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp. 412–419, Nov 2011.

[22] W. Wang, S. Yang, S. Bhardwaj, S. Vrudhula, F. Liu, and Y. Cao, "The impact of nbti effect on combinational circuit: Modeling, simulation, and analysis," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 18, pp. 173–183, Feb 2010.

[23] A. Stempkovsky, A. Glebov, and S. Gavrilov, "Calculation of stress probability for nbti-aware timing analysis," in *2009 10th International Symposium on Quality Electronic Design*, pp. 714–718, March 2009.

[24] L. Zhang, W. Chen, Y. Hu, and C. C.-p. Chen, "Statistical Timing Analysis with Extended Pseudo-Canonical Timing Model," *Design, Automation and Test in Europe*, pp. 952–957, 2005.

[25] C. E. Clark, "The greatest of a finite set of random variables," *Operations Research*, vol. 9, no. 2, pp. 145–162, 1961.

[26] C. Visweswariah, K. Ravindran, K. Kalafala, S. G. Walker, S. Narayan, D. K. Beece, J. Piaget, N. Venkateswaran, and J. G. Hemmett, "First-Order Incremental Block-Based Statistical Timing Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 10, pp. 2170–2180, 2006.

[27] http://www.nangate.com. Accessed 2018-10-24.

[28] A. Agarwal, D. Blaauw, V. Zolotov, S. Sundareswaran, M. Zhao, K. Gala, and R. Panda, "Statistical delay computation considering spatial correlations," in *Proceedings of the ASP-DAC Asia and South Pacific Design Automation Conference, 2003.*, pp. 271–276, Jan 2003.

[29] J. Chung, J. Xiong, V. Zolotov, and J. A. Abraham, "Path criticality computation in parameterized statistical timing analysis using a novel operator," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 4, pp. 497–508, 2012.

P. R. Chithira received the B.Tech. degree in electronics and communication engineering from Model Engineering College, Cochin, India, and the M.Tech degree in Microelectronics and VLSI design from National Institute of Technology, Calicut, India, in 2009 and 2011 respectively. She is currently pursuing the Ph.D degree with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India.

Her current research interests include statistical timing analysis and circuit reliability analysis.

Vinita Vasudevan received the B.Tech. degree in engineering physics from the Indian Institute of Technology Bombay, Mumbai, India, the M.S. degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, USA, and the Ph.D. degree in electrical engineering from the Indian Institute of Technology Bombay, in 1986, 1988, and 1993, respectively.

She is currently a Professor with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India. Her current research interests include CAD for analog and mixed signal circuits, statistical timing and power analysis and analysis of noise in circuits and systems.