IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 20, 2021, accepted August 4, 2021, date of publication August 9, 2021, date of current version August 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3103200

Lifetime Reliability Improvement of Nano-Scale Digital Circuits Using Dual Threshold Voltage Assignment

MOHSEN RAJI [1], REZA MAHMOUDI [1], BEHNAM GHAVAMI [2], (Member, IEEE), AND SAEED KESHAVARZI [1]
[1]School of Electrical and Computer Engineering, Shiraz University, Shiraz 71946-84636, Iran
[2]Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman 76169-14111, Iran
Corresponding author: Mohsen Raji (mraji@shirazu.ac.ir)

ABSTRACT In nano-scale CMOS technology, circuit reliability is a growing concern for complicated digital circuits due to manufacturing process variation and aging effects. In this paper, a statistical circuit optimization framework is presented to analyze and improve the lifetime reliability of digital circuits in the presence of process variation and aging degradation. The proposed framework takes advantage of a process variation- and aging-aware gate-level delay degradation model to characterize and evaluate the lifetime reliability of combinational circuits. A metric called Guardband-Aware Reliability (abbreviated as GAR) is proposed for a fair evaluation of the lifetime reliability of combinational circuits considering a guardband and timing yield specified by the designer. Then, using a criticality metric, a set of statistically critical gates is selected for being optimized in the optimization framework. As the improvement procedure, the dual-threshold voltage assignment technique is applied to the identified critical gates to enable the manufactured chip to improve the lifetime reliability in terms of low timing yield loss. Experimental results on ISCAS'85 and ISCAS'89 benchmark circuits show that our proposed framework increases the circuit reliability up to 9.93% for a 6-year lifetime imposing less than 6.9% timing yield loss.

INDEX TERMS Combinational circuits, dual threshold voltage assignment, optimization, lifetime reliability, process variations.

I. INTRODUCTION

Although technology scaling allows to fabricate chips with higher complexity and performance, it poses a severe challenge for reliable digital circuit design. With the scaling down of the transistor dimensions, fabrication-induced process variation (PV) causes the timing of the manufactured circuit to significantly deviates from its initial design [1]. As an example, with process technology scaling from 350 nm to 45 nm, the timing yields of the integrated circuits (ICs) have reduced from nearly 90% to mere 30% [2]. On the other hand, the aging mechanisms, such as Bias Temperature Instability (BTI), are appeared in nano-scale technology leading to considerable degradation of the conductance of MOS transistors during lifetime [3]–[6]. Negative BTI occurring in PMOS transistors and Positive BTI affecting NMOS transistors increases the absolute value of threshold voltage (Vth) of transistors and hence, the delay of the circuits increases with the operation time [7]. Consequently, it is necessary to consider the effects of PV and aging (specifically BTI mechanisms) in the analysis and improvement of lifetime reliability in nano-scale digital circuits.

There are many works which have analyzed the impacts of PV and BTI separately [2], [8]–[11]. However, considering the interdependencies between BTI and PV [13], [14], such separate analysis leads to inaccurate and unrealistic results. Recently, some papers have addressed the combined impacts of process variability and NBTI or BTI at different levels of abstraction such as device-level [9], gate-level [15]–[17], and architecture-level [18]–[20]. Siddiqua *et al.* [21] explored both NBTI and PV in an SRAM cell. In [13], an NBTI-aging aware delay degradation model for the logic gates in the presence of PV is proposed. Probability density functions (PDF) is used to represent the gate delays at manufacturing time. Then, the model presents a PDF for the gate delay at a given year of stress time. A statistical circuit

The associate editor coordinating the review of this manuscript and approving it for publication was Hao Luo.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

optimization flow introduced in [15] which improve combinational circuit lifetime reliability in the presence of the joint effect of PV and NBTI. Han and Kim [22] proposed a stochastic model of the Vth value variation and the gate aging delay time which considers the effects of the PV on NBTI. Perez-Rivera *et al.* [14] have analyzed the impact of PV and NBTI effect on IC lifetime reliability. Although the aforementioned work analyzed combined impacts of NBTI and PV, their analysis is limited to NBTI rather than covering both NBTI and PBTI while it is shown that PBTI is changing to a serious aging effect in modern VLSI systems [23], [24]. Some research have studied both NBTI and PBTI effects on circuit reliability, In [7], combinational circuits a reliability considering PVT dependence of NBTI and PBTI aging effects is analyzed. Duch *et al.* [16] present simulation based (HSPICE Monte Carlo) analysis of both NBTI and PBTI incorporate with parameter variations in logic gates. In [24], an analytical methodology for accurate modeling of the correlation between Process, Voltage and Temperature variations (PVT) and BTI-induced aging is presented. In [25], a machine learning approach is proposed to predict the NBTI degradation of the circuit paths. A fast, accurate, and versatile PV- and aging-aware delay model for generic cell libraries called AADAM is presented in [26]. Based on transistor-level SPICE simulations, the delay degradation of each library cell is characterized under a set of variability and aging factors. It is notable that PV- and aging analysis methods suffer from limitations: some of the previous works only consider NBTI effect (and neglect PBTI) [14], [15], [21]–[25] or some others just focused on one challenge (either PV or aging effect) [8]–[11] leading to inaccurate analysis results. Moreover, most of the other works, which consider both PV and aging effects, use the old reaction-diffusion theory to consider NBTI and PBTI effects on circuits [7], [16].

There are several approaches presented for mitigating the impact of either PV [11] or aging [27]–[31] in combinational circuits. Due to the interdependencies between PV and aging [14], [29], the mitigation technique which ignores PV or aging effects results in non-optimized solutions. Among the mitigation approaches in which both PV and aging effects are taken into consideration, gate sizing-based optimization techniques show their efficacy in the literature. In [23], mathematical methods (i.e. Lagrangian Relaxation) are used to optimize the circuit area considering circuit delay degradation. However, the computational complexity of these methods makes them impracticable for large-scale combinational circuits. Other approaches are based on upsizing a gate for delay degradation reduction in expense of area cost [14], [15], [17], [32]. However, most of these methods focus only on increasing the size of the selected gates to improve the delay of the critical aging paths leading to large area cost in the reliability optimized circuits. In [17], a method is proposed to reduce guardband by effective selection critical gates and using gate-sizing approach considering BTI and PV effects. This work focuses on finding potential critical path (PCP) (paths which are degraded more than other ones) to up/downsizing gate in path in term of low area overhead. However, since this work only uses gate sizing technique, it leads to a large area overhead to the VLSI designs. In [32], dynamic frequency scaling is proposed to mitigate the impacts of BTI-aging effects considering workload uncertainty and PV. In [12], a two-phase gate sizing approach is presented to improve the reliability of the circuit considering the joint effect of PV and transistor aging. In the first stage, the initial delay of the circuit is optimized to improve the timing yield of the circuit. Then, in the second stage, we reduce the delay degradation induced by aging and process variations. The previous lifetime reliability improvement approaches suffer from serious limitations; i.e. some of the previous works did not consider the joint effect of PV and aging leading to inefficient solutions while some others impose large area/delay overhead to the design. Hence, it is required to present a lifetime reliability with low design overhead and comprehensive analysis method considering the impacts of PV, NBTI, and PBTI.

In this paper, we present an optimization framework for maximizing circuit lifetime reliability in the presence of PV and aging effects using dual threshold voltage (DVth) technique. Firstly, we introduce a novel lifetime reliability evaluation metric called Guardband-Aware Reliability (abbreviated as GAR) metric to analyze the lifetime reliability of combinational circuits considering a guardband and timing yield specified by the designer. Also, we introduce a criticality metric to efficiently identify the best gates as the candidates for applying DVth assignment technique. Then, based on a statistical gate-level delay modeling, three optimization algorithms are proposed to improve the lifetime reliability of combinational circuits in the presence of PV and aging effects, as follows:

- Greedy-based reliability optimization algorithm (abbreviated as GeRO),
- Metaheuristic-based reliability optimization approach using simulated-annealing (SA) technique (abbreviated as SARO),
- Sensitivity based reliability improvement algorithm based on TILOS approach [34] (abbreviated as TIRO).

In these three optimizations algorithm, there is a tradeoff between optimization improvement and optimization speed. GeRO and SARO are fast while TIRO achieves more reliability improvements due to the proposed sensitivity metric. On average, the experimental results on ISCAS'85 and ISCAS'89 benchmarks circuits show that the proposed method increases the circuit reliability up to 6.38%, 8.16% and 9.93% imposing 6.50%, 7.03% and 6.9% timing yield loss for a 6-year lifetime and 10% PV for GeRO, SARO and TIRO, respectively.

Briefly the main contributions of this paper are as follows:

- DVth assignment technique is proposed for improving the robustness of combinational circuits against aging and PV effects,

- Guardband-Aware Reliability (abbreviated as GAR) metric is proposed to analyze the lifetime reliability of combinational circuits considering a guardband and timing yield specified by a designer i.e. it allows to designer to choose specific timing yield and optional guardband for calculating reliability,
- Three optimization algorithms (i.e. a GeRO, SARO and TIRO) for lifetime reliability improvement of combinational circuits.

The rest of this paper is organized as follows. Section 2 brings some backgrounds to the readers such as the reliability concepts in digital circuits and also, introduces the statistical gate delay degradation model which is used in this paper. In section 3, the dual threshold voltage assignment technique optimization is described while section 4 is about the critical gate identification approach. Section 5 describes the experimental results and finally, section 6 concludes the paper.

II. BACKGROUND

This section provides some necessary backgrounds on statistical static timing analysis and the impacts of aging on transistor delays.

A. STATISTICAL STATIC TIMING ANALYSIS

For timing analysis of digital circuits, it is required to compute the maximum delay of all paths from primary inputs to the primary outputs, irrespective to input signals. Such maximum delay is computed using a static simulation which is called static timing analysis (STA) [34]. STA has shown it efficacy in characterizing the timing performance of digital circuits, determining the delay information, and finding the critical paths. There are three important timing concepts as follows:

- Arrival Time (AT) which is defined as the latest time in which the final correct value of a signal arrives at a certain internal node (i.e. the input of a gate or cell). In computing AT, the signal transitions are propagated through the gates considering all the net and logic gate delays in between the reference input point and the destination node.
- Required Arrival Time (RAT) defined as the latest time at which a signal can arrive to a certain internal node without increasing the assumed delay of the circuit. RAT is computed by backward propagation from the inputs of the target storage elements (i.e. flip-flops) in the circuit to the internal nodes.
- Slack which is the difference between the RAT and the AT for each signal.

STA propagates actual ATs and RATs of the internal circuit nodes. In a circuit with correct timing behavior, the slack of all nodes is required to be larger or equal to zero.

In nano-scale circuit, Statistical STA (SSTA) allows to properly considering the impact of PV on path behavior and hence, a more accurate set of path delays. In SSTA, the net/gate delays are modeled as probabilistic random variables considering the effects of PV. A canonical first-order delay model was proposed in [34] considering the impact of PV as:

$$a_0 + \sum_{i=1}^{n} a_i \Delta X_i + a_{n+1} \Delta R_a \tag{1}$$

where a_0 is deterministic part, $\sum_{i=1}^{n} a_i \Delta X_i$ is correlated portion of PV effects and $a_{n+1} \Delta R_a$ is the independent portion. $\Delta X_i, i = 1 \ldots n$ shows the fluctuation on n global source of variation X_i computed by subtracting their mean value $\Delta X_i = X_i - \hat{X}_i$. The variables $a_i, i = 1 \ldots n$ are the gate delay sensitivity to each of the global variation source. ΔR_a is the variation of an independent random variable R_a from its nominal value while a_{n+1} is the sensitivity of the gate delay (or other timing quantities) to uncorrelated variations.

Based on the probabilistic delay models, SSTA estimates the probability distribution of the circuit performance under the variation of process parameter in a single timing analysis with a statistical approach. In SSTA, the mathematical operators (i.e. SUM and MAX) which are traditionally deterministic in STA are replaced with statistical ones [34].

B. BTI AGING MECHANISM

In nano-scale digital circuits, BTI is the main aging mechanism which affects MOSFET transistors. There are two types of BTI effects: Negative and Positive BTI (NBTI and PBTI) which respectively affect PMOS and NMOS transistors when these transistors work in linear or saturation regions. Preliminarily, NBTI was considered the significant reliability issue but after introducing the high-k metal gate dielectric in deep nano-scale (<45 nm) technologies [12], PBTI effects has also become an important aging issue. BTI mechanisms consist two phases [35]:

- *Stress Phase:* High vertical electric field, applied to transistor channel leads to breaking $S_i - H$ bonds and degrading device $S_i - S_iO_2$ interface. Two single H atom combine each other and generate H_2 molecule which leaves the so-called interface trap [34]. Also, charge trapping in thin silicon oxide layer is another reason of BTI. These phenomena gradually cause increasing threshold voltage (Vth) during lifetime.
- *Recovery Phase:* When transistor switched to cutoff region ($|V_{gs}| = 0$) some of the traps in the $S_i - S_iO_2$ interface is passivated. Therefore, it is observed some partial recovery in Vth degradation during the stress phase.

Maximum increasing in Vth depends on device stress time (i.e. the percentage of time during which the device is on), named as *stress probability (SP)*. BTI-induced Vth degradation is calculated as Eq. (3) [7],

$$\Delta V_{th,BTI} \approx K.t_{ox}.\sqrt{C_{ox}.\left(V_{gs} - V_{th0}\right)}.e^{\frac{E_{ox}}{E_0}}.e^{\frac{-E_a}{k.T}}.a^n.t^n \tag{2}$$

where n is the time exponent, t_{ox} is the effective oxide thickness, E_{ox} is the vertical electric field, T is the temperature in kelvin unit. k is the Boltzmann constant, C_{ox} is the oxide

capacitance per unit of area, V_{th0} is the initial threshold voltage value, E_a is activation energy. E_0 is constant, a is the percentage of the time the gate is under stress and K is technology dependent fitted constant, which can be different for NBTI and PBTI. Eq. (3) shows that, Vth shift depends on the initial Vth (V_{th0}). On the other hand, PV causes variance in V_{th0} in nano-scale digital circuits. Using the first-order Taylor approximation, the impact of PV in the long-term degradation of Vth is modeled as [15],

$$\Delta V_{th,BTI} = A.\left(1 - \gamma.\Delta V_{th.PV}\right).a^n.t^n \quad (3)$$

where $V_{th.PV}$ is the shift in V_{th0} due to PV, and A and γ are fitted constants. Then, the total Vth variation of a transistor m accords with the summation of the BTI ($\Delta V_{th,BTI}$) and PV ($\Delta V_{th.PV}$), as follows [15]:

$$\Delta V_{th,m} = A_m.a_m^n.t^n + \left(1 - \gamma \,.\, A_m \,.\, a_m^n \,.\, t^n\right).\beta_k.\sum_i \Delta V_{th(i)} \quad (4)$$

Note that, at the beginning of the device lifetime (i.e. $t = 0$ in Eq. (5)), the total variation of Vth is just due to PV effects. During device lifetime (i.e. with increasing t), BTI causes a shift in both the mean value and the variance of Vth [13].

C. PV- AND BTI- AWARE GATE DELAY MODEL

The PV- and NBTI/PBTI-aware delay of gate k (D_k) is considered as [15]:

$$D_k = D_{nom(k)} + B_k.a_k^n.t^n + \left(1 - \gamma.A.a_k^n.t^n\right).\beta_k.\sum_i \Delta V_{th(i)} \quad (5)$$

where $D_{nom(k)}$ is the nominal gate delay. B_k is a fitting parameter which shows increasing gate delay caused by BTI-induced threshold voltage increase under nominal condition. Also, β_k is a fitted coefficient which indicates the of PV-induced threshold voltage variation effects on the gate delay change without considering the BTI effect.

BTI-induced Vth increase depends on the fraction of time that the transistor is under stress (i.e., input signal is logic '0' for PMOS and '1' for NMOS) for a period of time. It is called as stress probability (SP). Note the difference between SP and the statistical signal probability which is typically defined as the probability that the input signal is logic '1'.

Due to the complementary behavior of pull-up and pull-down network of CMOS gates, the gate is in maximum stress when the signal probability. In order to consider the impact of both NBTI and PBTI, the a value is computed as:

$$a = |0.5 - logic\ gate\ inputs\ signal\ probablity| \quad (6)$$

Note that, Vth contains two deviation parts; the first part is associated to the time-zero variation and the other one is associated to aging effects (See Eq. (5)). The computational complexity of this model is low and the error which spread by discarding high-order terms in this linear model can be ignored.

Before using SSTA, parameters in Eq. (6) are pre-computed by using HSPICE simulations for basic gate type (i.e., INV, BUFF, NAND, NOR) at different design.

III. GUARDBAND AWARE LIFETIME RELIABILITY (GAR) METRIC

In this section, we propose a novel metric to analyze and evaluate the combinational circuit lifetime reliability called as Guardband Aware timing Reliability (GAR).[1]

In order to obtain GAR metric, we first define the concept of lifetime reliability for a combinational circuit. Based on the traditional formal definition of reliability, combinational circuit lifetime reliability at time t is defined as the probability of operating circuit at time t correctly given the circuit works correctly at time 0 (fresh time)); i.e.:

$$\mathcal{R}(t) = \mathcal{P}(operational\ at\ t \mid operational\ at\ 0) \quad (7)$$

A combinational circuit works properly at time 0 when the delay of the Critical Path (CP) is less than a given timing constraint (τ). A circuit is reliable at the end of a specific lifetime if the CP delay at that time (t) remains less than the same timing constraint at the fresh time (i.e. τ). Hence, Eq. (7) can be rewritten as:

$$\mathcal{R}(t) = \mathcal{P}(d_t < \tau \mid d_0 < \tau) \quad (8)$$

where d_t and d_0 respectively represent the values of circuit CP delay at time t and time 0, and τ is the timing constraint of the combinational circuit defined by the designer at the design stage.

In order to solve Eq. (8) to find the lifetime reliability, we use the conditional probability calculation rules. Hence, we have:

$$\mathcal{R}(t) = \frac{\mathcal{P}(d_t < \tau \cap d_0 < \tau)}{\mathcal{P}(d_0 < \tau)} \quad (9)$$

where $\mathcal{P}(d_t < \tau \cap d_0 < \tau)$ shows the probability of the intersection of two events $d_t < \tau$ and $d_0 < \tau$ (i.e. combinational circuit CP delay at time τ and time 0 be less than the timing constraint τ). Since the CP delay increases by increasing the operation time of the combinational circuit due to the BTI effects [3], we have:

$$d_0 < d_t \quad (10)$$

Therefore, if the event $d_t < \tau$ occurs, event $d_0 < \tau$ has certainly occurred. So, we have:

$$d_0 < \tau \subset d_t < \tau \quad (11)$$

where $\subset$ shows that event $d_0 < \tau$ is a subset of event $d_t < \tau$. Based on the rules of probability calculation, we have:

$$\mathcal{P}(d_t < \tau \cap d_0 < \tau) = \mathcal{P}(d_t < \tau) \quad (12)$$

Substituting Eq. (12) in Eq. (9), the lifetime reliability is obtained as:

$$\mathcal{R}(t) = \frac{\mathcal{P}(d_t < \tau)}{\mathcal{P}(d_0 < \tau)} \quad (13)$$

[1] The procedure of computing GAR is similar to our previously introduced timing reliability metric called TYR [53]. However, it is notable that, TYR was applicable for FFs and also, we extend TYR to a guardband-aware metric to consider the guardband which are assumed by the designers to make their designs reliable against long-term aging variation effects.

Due to the impacts of variabilities, combinational circuit CP delay is modeled as a random variable with a Gaussian distribution [39], [38]. So, $\mathcal{P}(d_0 < \tau)$ is equal to the value of Cumulative Distribution Function (CDF) of CP delay at time 0 for the timing constraint τ. Designers specify the variation-aware timing constraint for combinational circuits by using the concept of *p*-percentile point of CP delay CDF; i.e. the value for which the CDF is equal to *p* (Figure 1).

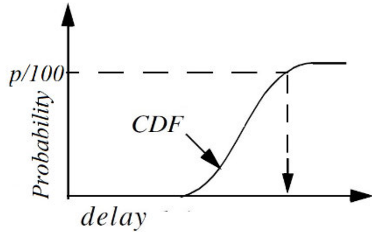

FIGURE 1. ***p*-percentile point concept.**

Designers of high reliable circuits consider a guardband for the obtained p-percentile point value of CP CDF to ensure the targeted lifetime reliability [17], [26]. So, for a reliability-aware circuit design, the timing constraint (τ) is computed as:

$$\tau = (1 + g) \times \varphi_0^{-1}(p/100) \tag{14}$$

where φ_0^{-1} shows the inverse CDF of CP delay at time 0 and g shows the guardband value ($0 \leq g \leq 1$) considered by the designer. Therefore, Eq. (13) can be re-expressed as:

$$\mathcal{R}_g^p(t) = \frac{\varphi_t\left((1+g) \times \varphi_0^{-1}(p/100)\right)}{\varphi_0((1+g) \times \varphi_0^{-1}(p/100))} \tag{15}$$

where $\mathcal{R}_g^p(t)$ is the GAR metric for the lifetime reliability of the combinational circuit at time t considering guardband value of g for the p-percentile point value of CP CDF, φ_t and φ_0 show the CDF of CP delay at time t and 0, respectively.

The GAR degradation for a specific guardband g and p-percentile point value at operational time t ($\Delta\mathcal{R}_g^p(t)$) is computed as:

$$\Delta\mathcal{R}_g^p(t) = \frac{\mathcal{R}_g^p(t) - \mathcal{R}_g^p(0)}{\mathcal{R}_g^p(0)} \tag{16}$$

IV. DUAL THRESHOLD VOLTAGE ASSIGNMENT TECHNIQUE

Dual threshold voltage assignment (DVth) is widely used for low-voltage low-power and low leakage power applications [27]. In dual Vth designs, higher Vth is assigned to the gates in non-critical paths for reducing the leakage power, and lower Vth is assigned to the gates in critical paths for improving performance/timing yield [27]. Implementing dual Vth technique is easy to fabricate by using an additional mask layer [27]. Here, we propose applying dual Vth assignment technique for reducing the impact of BTI on the circuit delay for improving the lifetime reliability. A transistor with higher Vth is expected to experience less BTI effect (Vth shift) due to the reduction in the electric field stress in the oxide (*Eox*) consign Eq. (3) [41], [42]. Here we provide a motivation example to show DVth effect on BTI-induced circuit delay degradation.

A. MOTIVATION EXAMPLE

Figure 2 shows C17 circuit of ISCAS'85 benchmark suite. We conducted a set of experiments in which three different Vths are assigned to all the gates in the circuit and each time, we measure the delay of the highlighted path (from input node N1 to output node N22. Figure 3 shows GAR values ($\mathcal{R}_{0.1}^{0.99}(t)$ for $0 \leq t \leq 10$) of C17 circuit with three different Vths. It is observed that, higher Vth leads to higher lifetime reliability. For example, the reliability of the circuit after the 10-year lifetime is around 0.86 when Vth of 660mV is assigned to the gates while it is reduced to 0.8 for Vth equal to 480mV.

B. DVth ASSIGNMENT TECHNIQUE OVERHEAD

Although, higher Vth results in less BTI-induced reliability degradation, using higher Vth imposes delay overhead to the design leading to timing yield loss in the circuit. Figure 4 shows the delay of the path considered in the motivation example during 10 years of circuit lifetime. It is observed that, the delay of the circuit is increased with the increase of the lifetime

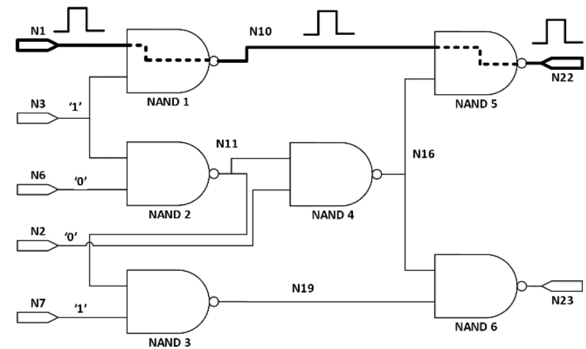

FIGURE 2. **C17 benchmark structure.**

In order to formally evaluate the cost of DVth assignment technique, we describe cost function ($Cost_g^p$) as below:

$$Cost_g^p = \frac{\varphi_L\left((1+g) \times \varphi_L^{-1}(p/100)\right) - \varphi_H\left((1+g) \times \varphi_L^{-1}(p/100)\right)}{\varphi_L\left((1+g) \times \varphi_L^{-1}(p/100)\right)} \tag{17}$$

where φ_L and φ_H show CDF of CP delay at time 0 for low and high Vth design, respectively, and φ_L^{-1} indicates the inverse CDF of CP delay at time 0 for low Vth circuit. The cost function reflects the timing yield loss due to the increased delay caused by DVth technique.

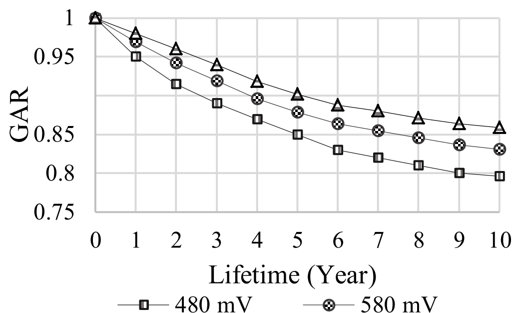

FIGURE 3. GAR value of highlighted path in figure 4 for different Vth value during 10 years lifetime.

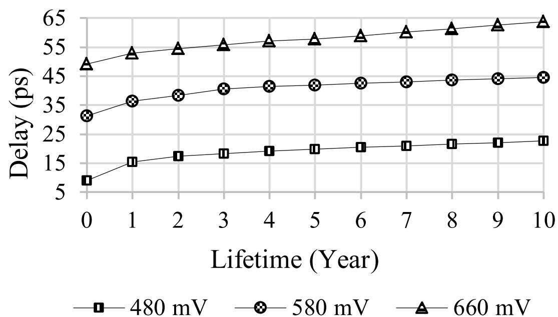

FIGURE 4. Delay of highlighted path for different Vth.

In order to investigate the overhead of DVth assignment technique, we provide an example in which DVth is applied to C17 in different cases. The obtained results of these experiments are presented in Table 1. The first column shows different cases of applying DVth assignment technique; i.e. "No Gate" refers to the case in which no gateis assigned with high Vth, "All gates in the path" indicates the case in which all the gates in the highlighted path in Fig 4 are assigned with high Vth and the other ones have low Vth, and "NAND 5" shows the case in which only gate NAND 5 is assigned with high Vth and the other ones have low Vth. The other columns show the initial delay of the circuit (99-percentile point of CP CDF), the cost ($Cost_{0.1}^{99}$ based on Eq. (17)), and GAR degradation ($\Delta\mathcal{R}_{0.1}^{99}$ (10) based on Eq. (16)). As the results show, using higher Vth results in lower GAR degradation but higher cost and higher initial delay. Moreover, it is observed that, when DVth assignment is used selectively (to only NAND 5 gate), the lifetime reliability degradation is reduced along with less imposed cost thus, providing reliability-cost tradeoff. Hence, if appropriate gates are selected for applying DVth assignment technique, the lifetime reliability of the circuit can be improved with controlling the imposed cost. On the other hand, for a combinational circuit with n gates, the number of cases of applying DVth assignment technique is 2^n circuit; i.e. each gate can be either assigned with low or high Vth. So, applying DVth assignment technique to combinational circuits is not a straightforward task. In the next section, we introduce an optimization flow and then, proposed three different algorithms to be used for lifetime reliability improvement of combinational circuits using DVth assignment technique.

TABLE 1. Cost and GAR degradation for different gate selection.

DVth Gate	Initial Delay (ps)	Cost (%)	$\Delta\mathcal{R}_{0.1}^{99}(10)$ (%)
No Gate	9.13	0	49
All gates in the path	31.38	73	11
NAND 5	17.45	8	14

V. LIFETIME RELIABILITY OPTIMIZATION FLOW

The overall view of the proposed lifetime reliability optimization flow is shown in Figure 5. Based on HSPICE simulations, BTI model, and PV effect information, the gate delay model is constructed and then, it is incorporated in the aging-aware SSTA considering duty cycle calculations. Based on SSTA results, a PV- and BTI-aware criticality metric is used to identify the statistically critical gate set as the candidates of applying DVth assignment technique. Then, DVth technique is applied to (one or more) critical gate(s) in the circuit based on the policy of the optimization algorithm; i.e. greedy-based method (GeRO), Simulated Annealing (SA)-based method (SARO), and a sensitivity-based (TIRO) method. Then, the critical gate set is updated and this flow continues until the critical gate set is empty. The algorithm always terminates as there is only two cases for each gate; i.e. it can be with either low Vth or high Vth.

In the following, we first present the PV- and BTI-aware criticality metric and then, three different approaches for lifetime reliability improvement are presented.

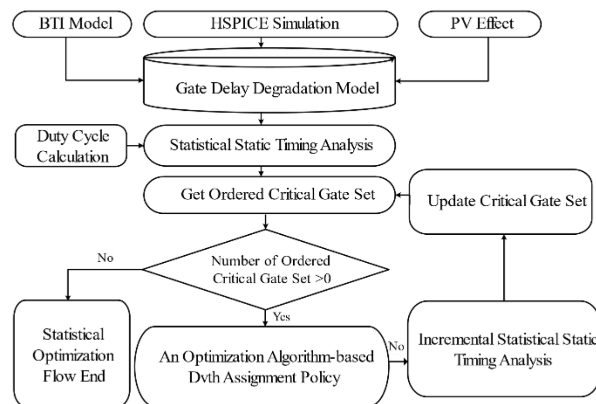

FIGURE 5. Overall flow of the proposed reliability optimization framework.

A. PROCESS VARIATION- AND BTI-AWARE CRITICALITY METRIC

In order to identify the candidate gates for applying DVth assignment technique, a PV- and BTI-aware criticality metric is proposed. The criticality of gate i is defined as:

$$Cririality\ (i) = \frac{\partial D_i}{\partial Vth_i} \times Slack\ (i) \tag{18}$$

where $\frac{\partial D_i}{\partial Vth_i}$ is Vth variation-induced gate delay deviation and $slack(i)$ is the gate delay slack and computed as:

$$Slack_{(Gate\ i)} = RAT_{(Gate\ i)} - AT_{(Gate\ i)} \tag{19}$$

1) OPTIMIZATION APPROACH #1: GREEDY-BASED METHOD (GERO)

A greedy algorithm tries to find local optimum choice at each stage with the intent of finding the global optimum [43]. Although a greedy strategy may not produce a global optimal solution, it achieves near-global optimum solutions in a reasonable amount of time [44]. Algorithm (1) shows the pseudocode of GeRO method for the circuit lifetime reliability optimization. After computation criticality of each gate, gates are ranked in descending order based on the computed criticality described in Eq. (18). Then, the most critical gate is picked out for assigning high Vth. Then, the circuit timing is computed using an incremental SSTA [31]; i.e. only the delay information of the gates in the fan-in cone of the modified gate is recalculated resulting in less computation time. In order to evaluate the impacts of DVth assignment technique, GAR metric described in Section IV is used. If the GAR degradation value of new circuit is less than the old one and the cost value computed based on Eq. (17) is less than the user-specific constraint, high Vth assignment to the current gate is considered as an acceptable move in the algorithm and the criticality list will be updated; otherwise, the circuit is roll-backed (i.e. the selected gate Vth is changed back to low Vth) and the gate is removed from the critical list. The process is repeated until the critical list becomes empty.

Algorithm 1 Greedy-Based Lifetime Reliability Optimization Method (GeRO)

Inputs: Netlist, Ordered candidate list (C_e), **Overhead constraint**
Output: Optimized Circuit

1. For each gate i in C_e List
   - 1.1. Assign high Vth to Gate i
   - 1.2. Recompute Timing of circuit incrementally
   - 1.3. If cost < user constraint and GAR degradation of the new circuit < GAR degradation of the previous circuit
     - 1.3.1. Accept new solution
     - 1.3.2. Terminate Optimization Policy
   - 1.4. Else
     - 1.4.1. Reject new solution and mark as an impossible candidate
2. End

2) OPTIMIZATION APPROACH #2: SIMULATED ANNEALING-BASED METHOD (SARO)

Simulated annealing (SA) is a metaheuristic approximate global optimization in large discrete search space from an optimization problem [45]. It is most suited for the problems in which finding an approximate global optimum is more important than finding a precise local optimum in a fixed amount of time.

The pseudo-code of the proposed SARO algorithm is presented in Algorithm (2). The algorithm starts by assigning low Vth to all gates in the circuit as the initial solution (x_c). Then, the criticality of all gates is computed using Eq. (18) and the most critical gate is picked out for high Vth assignment. Afterwards, the circuit timing is computed incrementally for the modified circuit which is called the new solution (x_n). If the new solution is better in terms of GAR degradation the new solution will be accepted and the current solution will be replaced by the new one. Otherwise, if the new solution is worse, it may be still accepted if a randomly generated number between 0 and 1 is higher than probabilistic value pr computed as:

$$pr = e^{\left(\frac{GAR\ deg(x_c) - GAR\ deg(x_n)}{temp}\right)} \tag{20}$$

where $GAR\ deg\ (x_c)$ and $GAR\ deg\ (x_n)$ are respectively the GAR degradation of solution (c) and new solution (n) computed using Eq. (16), and $temp$ is the temperature which is used to determine acceptance probability of worse solution. To avoid premature convergence, the rate of $temp$ reduction tends to zero [45] as:

$$T_{i+1} = \alpha * T_i \tag{21}$$

where i is the number of iterations and α is the cooling rate ($0 < \alpha < 1$).

Accepting worse new solutions occurs more at the beginning of the SARO algorithm (due to high value of $temp$) an it is for avoiding local optimum solutions [46]. However, the value of $temp$ is reduced in the next iterations (based on Eq. 21)) and thus, only the improved ones are allowed resulting in faster algorithm convergence. Further information on SARO can be found in the survey work in [46]. The SARO algorithm terminates until there is no critical gate in the candidate set. Since the critical list size is limited (total gates in circuits at maximum), it guarantees that SARO optimization algorithm will converge finally.

3) OPTIMIZATION APPROACH #3: SENSITIVITY-BASED METHOD (TIRO)

In this section, we present a sensitivity-based lifetime reliability optimization algorithm using DVth assignment technique. In the proposed sensitivity-based optimization algorithm, threshold voltage of gates is incrementally increased one-by-one to determine the threshold voltage assignment that provides the best reliability value. For each gate, it is important to consider the relative reliability improvement of a given gate to the imposed cost of the optimization technique. So, we propose a sensitivity metric to measure and analysis the impact of dual Vth of the gates on the reduction of variation effect of the whole circuit. At first, the critical set is obtained based on Eq. (18) and then, for each gate in the critical

Algorithm 2 Simulated Annealing-Based Lifetime Reliability Optimization Method (SARO)

Inputs: Netlist, Ordered candidate list (C_e), **Overhead constraint**
Output: Optimized Circuit

1. Set initial solution x_c and temperature
2. For each gate i in C_e List
   - 2.1. Assign high Vth to Gate i
   - 2.2. Recompute Timing of circuit incrementally (new solution x_n)
   - 2.3. If probability in Eq. (10) > random (0,1)
     - 2.3.1. Accept new solution
     - 2.3.2. Terminate Optimization Policy
     - 2.3.3.
   - 2.4. Else
     - 2.4.1. Reject new solution and mark as an impossible candidate
3. End

list, the timing information is updated using the aging-aware statistical timing analysis. Then, the best gate which provides the most benefit (best GAR degradation improvement) to the imposed cost (the lowest delay overhead), is selected based on the sensitivity metric computed as:

$$Sensitivity_{(gate\ i)} = \frac{GAR\ deg\ .\ imp\ .\ (i)}{cost\ (i)} \quad (22)$$

where $cost\ (i)$ is computed based on Eq. (17) when high Vth is assigned to gate i and $GAR\ deg\ .\ imp\ .\ (i)$ shows the GAR degradation improvement and computed as:

$$GAR\ deg\ .\ imp. = \frac{\Delta\mathcal{R}_g^p\,(t)_{Before} - \Delta\mathcal{R}_g^p\,(t)_{After}}{\Delta\mathcal{R}_g^p\,(t)_{Before}} \quad (23)$$

where $\Delta\mathcal{R}_g^p\,(t)_{Before}$ and $\Delta\mathcal{R}_g^p\,(t)_{After}$ show the reliability degradation of the circuit before and after applying DVth technique (assigning high Vth to a chosen gate), respectively.

After assigning high Vth to the gate with the highest sensitivity, the critical list is updated. The optimization loop repeats until the critical list becomes empty. The pseudocode sensitivity-based lifetime reliability optimization algorithm is presented in Algorithm. (3).

Algorithm 3 TIRO Method

Inputs: Netlist, Ordered candidate list (C_e), **Overhead constraint**
Output: Optimized Circuit

1. For each gate i in C_e List
   - 1.1 Compute Sensitivity of gate i (Eq. (22))
2. Sort Sensitivity List
3. Assign high Vth to the gate with the most sensitivity and accept this solution
4. End

VI. EXPERIMENTAL RESULTS

In this section, we conduct a set of the experiments to investigate different aspects of the proposed lifetime reliability improvement approaches. Firstly, we study the error of the proposed aging-aware delay model and then, the reliability of ISCAS'85 and ISCAS'89 benchmark suits are analyzed in presence of aging and PV. Then, the lifetime reliability optimization results for different optimization algorithm are investigated.

A. EXPERIMENTS SETUP

The proposed approaches are implemented in C++ and run on a windows machine with a core i7 quad-core Intel processor (4.6 GHz) and 32 GB RAM. The proposed approach was applied to ISCAS'85 and ISCAS'89 benchmark circuits. In this paper, the primitive gates (i.e. INV, BUFF, 2- to 4-input NAND, and 2- to 4-input NOR) are used in the netlist synthesis process. It is notable that, ISCAS'89 benchmark circuits are sequential circuit i.e. they have a combinational part and storage element (such as Flip Flop (FF)). So, these circuits are converted to the combinational ones by removing FFs and adding the FF' inputs/outputs to the primary outputs (POs)/primary inputs (PIs) of the circuit. Table 2 brings information about benchmark circuits used in the following experiments. The first and second columns show the circuit name and the number of primitive gates in benchmarks. The third column shows the number of FFs in ISCAS'89 benchmark. The dash line in some rows indicate that, ISCAS'85 benchmark circuits have no FF. The last column presents the number of PIs and POs.

The value of fitting parameters A, B, γ, and β in the statistical aging model are computed using HSPICE simulations and MOSFET Model Reliability Analysis (MOSRA) [49]. The HSPICE simulations are accomplished under PTM 22nm technology model [50], supply voltage 1.1V and 354K temperature.

To model the spatial correlation, a 3-level quad-three partition is employed [34]. All the gates are distributed randomly in a 4×4 grid in the bottom level. Then, the random variables of the gate delays are computed depending on the hierarchy of the gate. Based on [51], the random variables in same level have the same probability distribution. In this work, we use different total PV effect on Vth, i.e. 5%, 7% and 10%. For 10% PV, it is split into two parts; i.e. the first part for considering the systematic variance of 6% and the other one for random variance of 8%.

Although assigning higher threshold voltage to the circuit gates decreases the power consumption [25], we did consider the improvements achieved by the proposed DVth on power consumption. The reason behind this is that, the goal of the proposed framework is to improve the circuit lifetime reliability (and not the power consumption). However, the proposed

DVth increases the circuit delay and hence, we consider the circuit timing yield as the cost in the experiments.

TABLE 2. ISCAS'85 and ISCAS'89 benchmark circuits information.

Bench.	# Gates	# Flip Flops	(#PI, #PO)
C880	383	-	(60,26)
C2670	1193	-	(233,140)
C5315	2307	-	(178,123)
C7552	3512	-	(207,108)
S420	196	16 D-FF	(19,2)
S820	289	5 D-FF	(18,19)
S1488	653	6 D-FF	(8,19)
S5378	2779	179 D-FF	(35,49)
S15850	9772	534 D-FF	(77,150)

B. PV AND BTI-AWARE DELAY DEGRADATION MODEL VERIFICATION

In order to verify the accuracy of the proposed PV- and BTI-aware gate-level delay degradation model, gate delay values obtained from Eq. (6) is compared with the ones obtained from Monte-Carlo based HSPICE (MCH) simulations under the same variation distribution and working conditions. Table 3 presents the relative error values between the proposed delay degradation model and MCH simulation for the primitive gates. As the results show, the maximum error on μ and σ values of gate delay between the proposed model and MCH simulation, is less than 4%. It is notable that, the error is originated from the parameter fitting process. Nevertheless, the verification results indicate that, the proposed variation- and BTI-aware delay model has acceptable accuracy to be used in a reliability improvement framework under PV and BTI effects.

TABLE 3. Error of PV- and BTI-aware gate delay model.

Error (%)	Inv	NAND			NOR		
		2	3	4	2	3	4
σ	1.02	1.82	2.07	3.11	1.46	2.48	3.47
μ	1.36	2.39	2.61	3.42	3.22	3.74	3.98

C. LIFETIME RELIABILITY ANALYSIS

In order to investigate the severity of the lifetime reliability challenge in nano-scale digital circuit, the proposed statistical PV- and BTI-degradation model is incorporated in an SSTA to calculate μ and σ of the circuit maximum arrival times. The analysis is performed considering various variation ratios and different operational time. Duty cycles of the signals are computed considering their signal probabilities. SP values of all primary inputs' (PIs) are set to 0.5 and for internal nodes, SP is calculated using the approach presented in [52]. All of the gates in the circuit are assigned initially with a low Vth.

Figure 6 shows the CP delay values (computed as $\mu + 3\sigma$ from the CP delay distribution) after 3, 6, and 9 years of operation time normalized to design CP delay. As expected, the delay values are increased (about 2× in most benchmark circuits after 9 years) due to the BTI effects by increasing the circuit operation time. So, BTI poses a serious challenge for satisfying the timing constraints of digital circuit in the operational lifetime.

The impacts of PV on lifetime reliability of digital circuits are also investigated. Figure 7 shows GAR in considering variation ratio of 5%, 7%, and 10%. As shown in the figure, by increasing PV amount (expected in the future technology nodes), GAR is also decreased indicating that, lifetime reliability improvement should be addressed considering the impacts of PV on digital circuits.

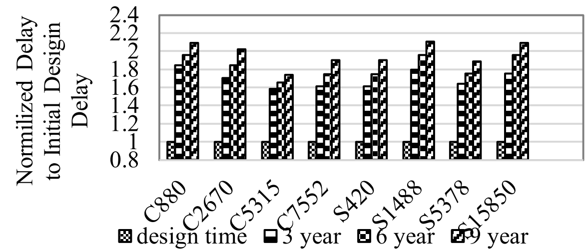

FIGURE 6. Normalized CP delay during lifetime.

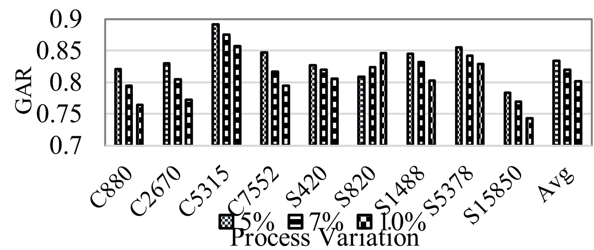

FIGURE 7. Impact of different PV value on GAR.

D. LIFETIME RELIABILITY OPTIMIZATION

In order to show the efficacy of the proposed lifetime reliability improvement approaches (GeRO, SARO and TIRO methods), the reliability of ISCAS'85 and ISCAS'89 benchmark circuits are evaluated under different variation ratios, operational lifetime, and guardband values.

At first, GAR degradation of the circuits under BTI effects (considering 3-, 6-, and 9-year operation time) and PV effects (considering 5%, 7%, and 10% variation) before applying any optimization technique is presented in Table 4. As the results show, GAR degradation is increased during the lifetime; for example, GAR degradation of C880 benchmark circuit is respectively 13.82%, 17.93% and 23.68% considering 3, 6, and 9-year operation time under 5% PV. Also, increasing PV effects leads to an increasing in GAR degradation; for example, GAR degradation of C880 benchmark considering 5%, 7% and 10% variation ratios are respectively 13.82%, 15.49 and 18.82 for 3-year lifetime.

Table 5-7, respectively show the GAR degradation values obtained from GeRO, SARO, and TIRO reliability improvement algorithms under BTI effects (considering 3, 6, and 9 years of operational time) and PV effects (considering 5%,

TABLE 4. Initial GAR degradation of the benchmark circuits (before applying any optimization method).

Bench.	Variation Ratio	5%			7%			10%		
	Operational Lifetime (y)	3	6	9	3	6	9	3	6	9
C880		13.82	17.93	23.68	15.49	20.57	26.84	18.82	23.59	29.30
C2670		12.01	17.02	23.08	14.51	19.55	25.13	16.41	22.80	29.73
C5315		7.58	10.79	13.84	9.29	12.46	15.35	11.91	14.27	17.81
C7552		9.62	15.26	20.83	12.18	18.29	23.21	14.21	20.55	24.86
S420		10.32	17.35	26.93	10.93	18.02	26.78	12.36	19.48	28.39
S820		9.85	14.90	19.22	11.02	17.65	22.97	13.16	19.08	25.42
S1488		10.84	15.47	21.06	11.90	16.76	22.56	14.46	19.74	25.32
S5378		6.95	9.90	12.64	10.33	15.81	19.78	12.01	17.12	21.77
S15850		15.41	21.72	28.52	15.75	23.14	29.19	18.34	25.72	31.96

TABLE 5. GAR degradation optimization results obtained by GeRO algorithm.

Bench.	Variation Ratio	5%			7%			10%		
	Operational Lifetime (y)	3	6	9	3	6	9	3	6	9
C880		8.23	12.69	14.66	9.49	14.03	16.63	12.57	17.00	22.66
C2670		7.91	11.85	16.49	9.64	14.25	18.22	11.34	16.04	22.51
C5315		4.82	6.29	8.15	5.91	7.63	9.47	7.07	8.98	11.32
C7552		6.15	9.03	10.99	7.77	11.18	13.02	9.38	12.88	14.50
S420		4.19	7.77	11.56	5.57	10.02	14.96	7.60	12.11	18.97
S820		7.84	10.63	15.48	9.56	12.74	17.12	9.85	14.90	19.20
S1488		7.43	10.20	13.61	8.88	11.38	15.36	10.47	14.35	18.36
S5378		4.25	6.38	9.04	5.17	7.61	10.87	6.95	9.90	12.64
S15850		7.47	12.99	16.90	8.44	14.21	18.28	10.80	16.35	21.90

TABLE 6. GAR degradation optimization results obtained by SARO algorithm.

Bench.	Variation Ratio	5%			7%			10%		
	Operational Lifetime (y)	3	6	9	3	6	9	3	6	9
C880		7.14	10.90	13.86	8.34	12.11	15.76	11.64	15.79	19.69
C2670		5.77	9.64	13.97	7.07	11.38	16.64	8.61	13.72	19.68
C5315		6.00	7.65	10.09	7.12	8.83	11.38	8.79	10.06	13.02
C7552		5.67	7.41	9.25	6.49	9.71	11.97	8.55	11.85	13.33
S420		4.19	7.77	11.56	5.57	10.02	14.96	7.60	12.11	18.97
S820		4.68	7.28	10.93	5.78	8.67	12.18	7.31	10.72	14.54
S1488		5.08	9.22	12.47	6.24	10.83	14.22	7.42	12.30	15.54
S5378		4.60	7.89	10.68	5.67	8.96	11.27	7.08	10.10	13.14
S15850		5.79	9.13	13.74	6.65	10.39	14.46	8.06	12.31	15.53

7%, and 10% variation) with 10% guardband. As shown in the tables, the proposed DVth technique incorporated in different optimization approaches leads to lifetime reliability improvement of digital circuits as the operational lifetime and PV effects are increased. For example, in GeRO approach, GAR degradation of C880 benchmark, for a 6-year lifetime and 5% PV is improved from 17.93% to 12.69% after optimization. Also, GAR degradations using SARO and TIRO optimization methods are reduced to 10.90% and 7.77%, respectively. In the most case, TIRO has the best improvement compared to GeRO and SARO methods. GeRO and SARO assign High Vth to the most critical gate which is selected based on the criticality metric considering the delay slack in addition to Vth sensitivity. So, a gate with a big slack and a medium Vth sensitivity may become the most critical gate in GeRO and SARO methods and thus, be selected as the most critical gate. Hence, assigning High Vth to the most critical gate in GeRO and SARO methods does not necessarily results in achieving the most reliability improvement. On the other hand, TIRO approach computes the sensitivity metric which is the relative reliability improvements per delay overhead achieved by assigning High Vth to each critical gate (i.e. in each optimization iteration of TIRO, all critical gates are analyzed and the gate with the most reliability improvement and the least delay overhead is assigned with High Vth). So, TIRO directly considers the reliability improvement and finds the gates for applying High Vth more intelligently and efficiently.

TABLE 7. GAR degradation optimization results obtained by TIRO algorithm.

Bench.	Variation Ratio	5%			7%			10%		
	Operational Lifetime (y)	3	6	9	3	6	9	3	6	9
C880		4.89	7.77	10.14	6.33	9.38	11.49	8.21	11.95	14.30
C2670		4.50	6.49	10.71	5.98	8.12	12.73	7.30	10.57	16.96
C5315		3.12	5.49	6.51	4.08	6.32	7.76	5.22	7.44	9.96
C7552		4.17	6.78	8.45	5.30	8.30	10.20	6.57	9.59	11.93
S420		4.19	7.77	11.56	5.57	10.02	14.96	7.60	12.11	18.97
S820		3.51	6.46	9.68	4.79	7.75	11.19	6.19	8.51	13.52
S1488		4.06	6.04	8.37	5.28	7.57	10.24	6.32	9.08	12.10
S5378		2.12	4.20	7.27	3.09	5.13	8.34	4.96	7.10	10.12
S15850		3.08	4.90	6.39	4.04	6.02	7.65	5.08	7.10	8.93

TABLE 8. Reliability improvement of GeRO, SARO, and TIRO.

Bench	GeRO		SARO		TIRO	
	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$
C880	5.06	7.47	5.06	7.47	5.06	7.47
C2670	7.47	8.15	8.84	7.93	13.06	5.42
C5315	6.17	5.58	4.90	5.22	7.96	6.92
C7552	5.62	5.78	6.59	5.06	8.69	5.09
S420	5.43	7.43	5.43	7.43	5.43	7.43
S820	5.16	5.91	10.33	8.53	11.41	5.85
S1488	4.66	6.91	8.23	8.23	9.50	9.50
S5378	8.69	5.48	8.45	5.15	12.08	8.18
S15850	9.23	5.87	12.58	8.32	16.23	6.26
Avg	6.38	6.50	8.16	7.03	9.93	6.90

TABLE 9. Reliability improvement and cost comparison of similar works and TIRO (3 years of lifetime, 5% variation ratio).

Bench.	[15]		[17]		TIRO	
	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$
C880	3.42	11.76	3.66	8.28	3.95	5.76
C2670	7.05	9.16	8.25	6.98	9.75	4.36
C5315	4.75	11.28	4.05	8.08	5.97	5.36
C7552	4.91	9.44	4.45	7.48	6.51	4.02
S420	3.25	10.78	4.25	8.36	4.25	5.44
S820	6.75	10.09	6.25	7.52	8.75	4.68
S1488	5.66	14.96	5.25	11.04	7.25	7.61
S5378	7.15	12.42	8.19	9.78	9.86	6.54
S15850	9.25	11.08	9.75	8.18	12.15	5.01
Avg.	5.79	11.21	6.01	8.41	7.61	5.42

In Table 8, the lifetime reliability improvement and the imposed timing yield overhead (i.e. Eq (17)) of the optimization approaches (i.e. GeRO, SARO and TIRO algorithms) for different benchmarks are reported under 10% variation ratio, 6 years of operational lifetime, and 10% guardband. The reliability improvement (GARINC) is computed as:

$$GARINC\ (\%) = \frac{GAR_{improved} - GAR_{initial}}{GAR_{initial}} \times 100 \quad (24)$$

where $GAR_{initial}$ and $GAR_{improved}$ respectively show the GAR values of the circuit before and after applying the optimization approach.

As the results show, on average, TIRO algorithm achieves 9.66% reliability improvement in expense of 6.86% cost while GeRO and SARO methods respectively achieve 6.24% and 7.78% reliability improvement by imposing 6.93% and 7.40% cost to the design. This shows that, TIRO improves circuit reliability more than GeRO and SARO with less delay cost comparing to GeRO and SARO. TIRO optimization considers improvement and cost simultaneously; i.e. TIRO chooses a gate to assign with High Vth which has the most GAR improvement per delay cost. So, it is expected that, TIRO results in more GARINC and less cost compared to the other two algorithms.

We also compare the three optimization algorithms (GeRO, SARO and TIRO) under different timing yield overheads ($Cost_g^p$ as descried in Eq (17)) (i.e. 5%, 10%, and 15%) with specific condition of BTI and PV effects (i.e. 6 years of

TABLE 10. Reliability improvement and cost comparison of similar works and TIRO (9 years of lifetime, 5% variation ratio).

Bench.	[15]		[17]		TIRO	
	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$
C880	3.62	13.24	4.06	9.22	4.45	6.24
C2670	7.55	10.34	9.05	8.27	10.25	5.14
C5315	5.25	12.72	4.55	9.02	6.67	6.64
C7552	5.01	10.56	4.95	8.62	7.61	4.23
S420	3.75	12.97	4.75	9.61	4.75	6.56
S820	7.25	11.35	6.75	8.48	9.62	5.82
S1488	6.26	17.04	5.75	12.96	7.75	8.15
S5378	7.65	14.83	9.09	11.47	10.46	7.21
S15850	10.75	12.42	10.25	9.07	13.65	5.15
Avg.	6.34	12.83	6.57	9.63	8.35	6.12

TABLE 11. Reliability improvement and cost comparison of similar works and TIRO (3 years of lifetime, 10% variation ratio).

Bench.	[15]		[17]		TIRO	
	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$
C880	4.14	13.88	4.52	9.70	4.81	6.98
C2670	9.14	11.38	10.26	8.14	11.90	5.38
C5315	5.68	14.04	5.52	9.64	7.44	6.80
C7552	5.52	11.50	5.66	9.34	8.28	4.76
S420	4.22	13.24	5.18	10.06	5.18	7.02
S820	8.06	12.76	8.16	8.96	10.64	5.88
S1488	6.71	17.68	6.86	13.64	8.80	8.96
S5378	9.02	15.12	10.18	11.54	11.42	7.72
S15850	11.96	13.84	12.06	9.74	15.18	5.86
Avg.	7.16	13.71	7.6	10.08	9.29	6.59

lifetime, 10% variation ratio). Figure 8 shows GAR degradation considering different $Cost_g^p$ obtained with different optimization algorithms. It is observed that, increasing timing yield overhead results in decreasing GAR degradation. That is because more freedom is provided for the optimization algorithms to assign more gates with High Vth leading to less GAR degradation (i.e. reliability improvement).

In order to compare the optimization procedure of the proposed algorithms, the steps of GAR degradation obtained in GeRO, SARO and TIRO optimization algorithms for S820 benchmark circuit are shown in Figure 9. Since GeRO approach always accepts a solution with better reliability than the previous one, its corresponding curve (the obtained GAR degradation) is monotonically decreasing. However, the SARO algorithm accepts bad solutions with a decreasing probability (p) and hence, its curve may be increasing in some points in the first steps of the optimization but will be decreasing in the next steps. TIRO approach accepts the best solution with the highest sensitivity value for each gate in the critical list in each iteration. Hence, its curve is also monotonically decreasing during the optimization steps. It is notable that, TIRO improves the circuit reliability faster than GeRO and SARO because TIRO directly chooses a gate with the highest sensitivity for assigning High Vth resulting in reducing GAR degradation in all steps. For example, in the first step of GeRO and SARO, the initial GAR degradation is decreased from 19.08% to 18.96% but TIRO chooses a different gate reducing the initial GAR degradation to 18.2%.

E. LIFETIME RELIABILITY OPTIMIZATION COMPARISONS

In order to show the efficacy of the proposed framework, we compare TIRO with circuit-level lifetime reliability optimization methods proposed in [15] and [17]. It is notable that, since there is no similar framework based on DVth technique, we choose these gate-sizing-based works since they take accounts the impacts of PV and BTI during lifetime reliability optimization.

Table 9-12 show the obtained comparison results in which the methods are compared in terms of GARINC (Eq (23)) and $Cost_{0.1}^{0.1}$ (Eq (17)) with specific condition of BTI and PV effect; i.e. 3 and 9 years of lifetime under 5% variation ratio in Table 9 and 10 respectively in addition to 3 and 9 years of lifetime under 10% variation ratio respectively in Tables 11 and 12. The experimental results show that TIRO dominates previous similar approaches in terms of both reliability improvement and imposed cost. For example, 9 years of lifetime and 10% variation ratio, TIRO averagely improves the lifetime reliability by 9.93% while the similar methods proposed in [15] and [17] respectively achieve 7.74% and 8.42% improvements in reliability. Also, TIRO imposes 6.90% cost while [15] and [17] lead to 14.51% and 10.42% cost during the circuit design, respectively. It is also notable that, one of the main superiority of our DVth-based technique over all previous circuit-level gate-sizing–based reliability improvement methods (including [15] and [17]) is that, our technique does not impose any area overhead to the design.

TABLE 12. Reliability improvement and cost comparison of similar works and TIRO (9 years of lifetime, 10% variation ratio).

Bench.	[15]		[17]		TIRO	
	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$	GARINC	$Cost_{0.1}^{0.1}$
C880	4.51	14.22	4.88	10.16	5.06	7.47
C2670	10.29	12.02	11.11	8.11	13.06	5.42
C5315	5.93	14.91	6.17	10.01	7.96	6.92
C7552	5.88	12.18	6.06	9.81	8.69	5.09
S420	4.70	13.36	5.43	10.42	5.43	7.43
S820	8.61	13.66	9.11	9.19	11.41	5.85
S1488	7.01	18.12	7.51	14.05	9.50	9.50
S5378	10.02	15.40	10.92	11.61	12.08	8.18
S15850	12.71	14.76	13.21	10.16	16.23	6.26
Avg.	7.74	14.51	8.27	10.42	9.93	6.90

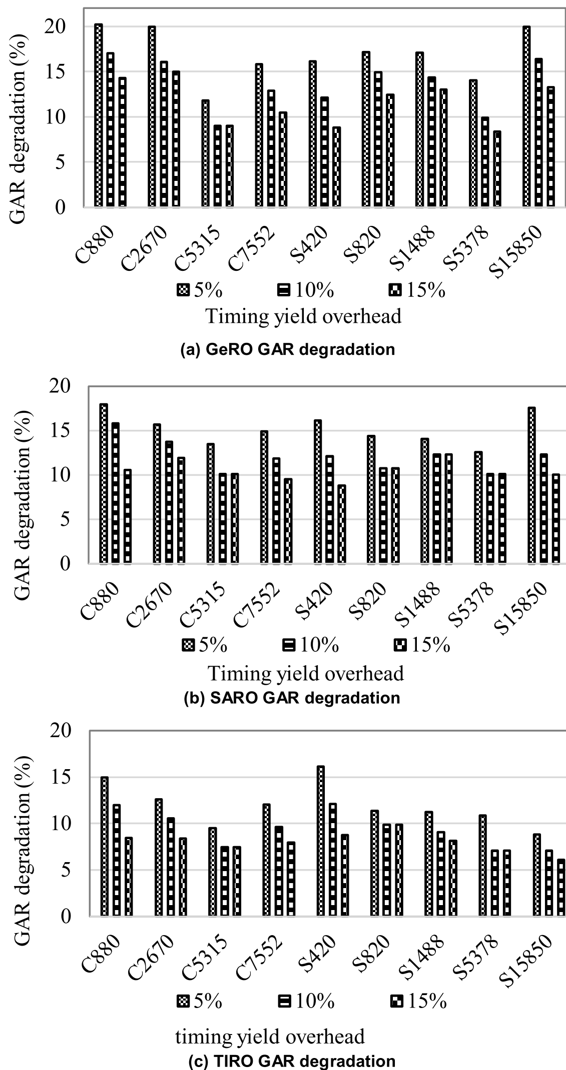

FIGURE 8. GAR degradation with different guardband values obtained from different optimization algorithms.

F. ALGORITHMS COMPUTATION COMPLEXITY AND RUNTIME

An important aspect of an optimization algorithm in computer-aided design is its computation complexity and runtime. Computation complexity of GeRO and SARO is $O(n)$ where n is the total number of gates in the combinational circuit. Because GeRO and SARO accept or reject the current critical gate for applying DVth technique, this gate will not appear in the next critical gate list. When the list becomes empty, the optimization will be terminated and thus, the optimization will be terminated at $O(n)$. On the other hand, TIRO firstly computes the reliability improvement per cost which is achieved by assigning High Vth to each gate in critical list. So, it is required to analyze the reliability and technique cost for all the gates in critical gates; i.e. the circuit is analyzed for all the critical gates in each iteration. As a result, TIRO has $O(n^2)$ time complexity. However, the results show that the number of critical gates is very much less than total number of gates in the circuits.

TABLE 13. Runtime comparison of the proposed optimization algorithms.

Bench.	Gates (PI,PO)	Runtime (s)		
		GeRO	SARO	TIRO
C880	(60,26)	10.4	10.6	100.9
C2670	(233,140)	351.2	342.3	1274.2
C5315	(178,123)	864	845	2709
C7552	(207,108)	1280	1211	3475
S420	(19,2)	1	1.2	9.2
S1488	(8,19)	12.4	9	63
S5378	(35,49)	990	964	2177
S15850	(150,77)	11990	12543	46532
Average	-	1626	1665	5834

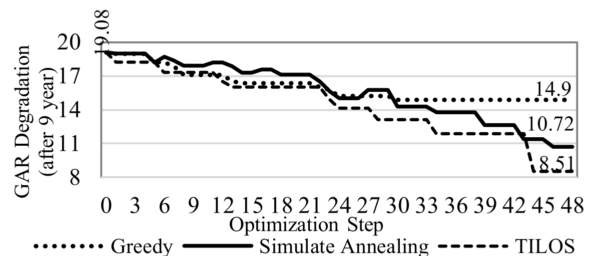

FIGURE 9. GeRO, SARO and TIRO optimization steps.

Table 13 shows the runtime of the proposed optimization algorithms for different benchmark circuits. The first two columns show the name and the information (number of gates, PIs and POs) of the benchmark circuits and the other three columns respectively represent the runtime of GeRO, SARO, and TIRO methods. It is observed that, the runtime

of GeRO and SARO methods are approximately equal (on average, GRO takes 1626 seconds and SARO takes 1665 seconds to finish) while TIRO has 5834 seconds runtime. As can be seen, GeRO and SARO have similar runtime while TIRO takes longer to optimize the circuit.

VII. CONCLUSION

Lifetime reliability of nano-scale digital circuits has encountered major challenges due to PV and BTI effects. This paper presents a statistical circuit optimization framework to optimize the circuit lifetime reliability under the joint effect of PV and BTI. Based on a novel metric for lifetime reliability evaluation, the circuit reliability is improved by assigning high Vth to the candidate gates which results in reducing BTI effects while increasing initial delay at the design time, i.e. timing yield overhead. The reliability improvement procedure is developed using three different optimization algorithms, i.e. greedy-based (GeRO), SA-based (SARO) and TILOS-like sensitivity-based (TIRO). The experimental results show averagely 6.38%, 8.16%, and 9.93% reliability improvement for GeRO, SARO, and TIRO optimization algorithms, respectively, while imposing 6.50%, 7.03% and 6.90% timing yield overhead, respectively. The obtained results show that, TILOS achieves the best reliability improvement with higher computation time.

There are many avenues as the future work of this paper. First, combining techniques such as gate resizing with dual Vth can be beneficial in achieving more reliability improvement as gate resizing decreases DVth timing yield overhead. Moreover, other evolutionary optimization algorithms (such as genetic algorithm) can be applied to gain higher circuit reliability improvements for combinational circuits.

REFERENCES

[1] K. A. Bowman, S. G. Duvall, and J. D. Meindl, ''Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration,'' *IEEE J. Solid-State Circuits*, vol. 37, no. 2, pp. 183–190, Feb. 2002.

[2] A. Agarwal, B. C. Paul, H. Mahmoodi, A. Datta, and K. Roy, ''A process-tolerant cache architecture for improved yield in nanoscale technologies,'' *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 1, pp. 27–37, Jan. 2005.

[3] V. M. van Santen, H. Amrouch, and J. Henkel, ''Modeling and mitigating time-dependent variability from the physical level to the circuit level,'' *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 66, no. 7, pp. 2671–2684, Jul. 2019, doi: 10.1109/TCSI.2019.2898006.

[4] J. Keane, X. Wang, D. Persaud, and C. H. Kim, ''An all-in-one silicon odometer for separately monitoring HCI, BTI, and TDDB,'' *IEEE J. Solid-State Circuits*, vol. 45, no. 4, pp. 817–829, Apr. 2010.

[5] B. Velamala, K. B. Sutaria, H. Shimizu, H. Awano, T. Sato, G. Wirth, and Y. Cao, ''Compact modeling of statistical BTI under trapping/detrapping,'' *IEEE Trans. Electron Devices*, vol. 60, no. 11, pp. 3645–3654, Nov. 2013.

[6] M. Grossi and M. Omaña, ''Impact of bias temperature instability (BTI) aging phenomenon on clock deskew buffers,'' *J. Electron. Test.*, vol. 35, no. 2, pp. 261–267, Apr. 2019.

[7] S. Krishnappa, H. Singh, and H. Mahmoodi, ''Incorporating effects of process, voltage, and temperature variation in BTI model for circuit design,'' in *Proc. IEEE Latin Amer. Symp. Circuits Syst.*, Feb. 2010, pp. 236–239.

[8] M. S. S. M. Saofi, H. Hussin, M. Muhamad, and Y. A. Wahab, ''Investigation of the NBTI and PBTI effects on multiplexer circuit performances,'' in *Proc. IEEE Int. Conf. Semiconductor Electron. (ICSE)*, Jul. 2020, pp. 49–52, doi: 10.1109/ICSE49846.2020.9166861.

[9] J. Fang and S. S. Sapatnekar, ''Understanding the impact of transistor-level BTI variability,'' in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Apr. 2012, Art. no. 6241887.

[10] N. Parihar, N. Goel, S. Mukhopadhyay, and S. Mahapatra, ''BTI analysis tool-modeling of NBTI DC, AC stress and recovery time kinetics, nitrogen impact, and EOL estimation,'' *IEEE Trans. Electron Devices*, vol. 65, no. 2, pp. 392–403, Feb. 2018.

[11] A. P. Shah and P. Girard, ''Impact of aging on soft error susceptibility in CMOS circuits,'' in *Proc. IEEE 26th Int. Symp. Line Test. Robust Syst. Design (IOLTS)*, Jul. 2020, pp. 1–4, doi: 10.1109/IOLTS50870.2020.9159733.

[12] S. M. Ebrahimipour, B. Ghavami, and M. Raji, ''A statistical gate sizing method for timing yield and lifetime reliability optimization of integrated circuits,'' *IEEE Trans. Emerg. Topics Comput.*, vol. 9, no. 2, pp. 759–773, Apr. 2021, doi: 10.1109/TETC.2020.2987946.

[13] W. Wang, V. Reddy, B. Yang, V. Balakrishnan, S. Krishnan, and Y. Cao, ''Statistical prediction of circuit aging under process variations,'' in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2008, pp. 13–16.

[14] Z. Perez-Rivera, E. Tlelo-Cuautle, and V. Champac, ''Gate sizing methodology with a novel accurate metric to improve circuit timing performance under process variations,'' *Technologies*, vol. 8, no. 25, pp. 1–12, 2020.

[15] S. Jin, Y. Han, H. Li, and X. Li, ''Statistical lifetime reliability optimization considering joint effect of process variation and aging,'' *Integration*, vol. 44, no. 3, pp. 185–191, Jun. 2011.

[16] L. Duch, M. Peon-Quiros, P. Weckx, A. Levisse, R. Braojos, F. Catthoor, and D. Atienza, ''Analysis of functional errors produced by long-term workload-dependent BTI degradation in ultralow power processors,'' *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 28, no. 10, pp. 2122–2133, Oct. 2020.

[17] A. Gomez and V. Champac, ''An efficient metric-guided gate sizing methodology for guardband reduction under process variations and aging effects,'' *J. Electron. Test.*, vol. 35, no. 1, pp. 87–100, Feb. 2019.

[18] H. M. Abbas, B. Halak, and M. Zwolinski, ''Learning-based BTI stress estimation and mitigation in multi-core processor systems,'' *Microprocessors Microsyst.*, vol. 81, Mar. 2021, Art. no. 103713.

[19] I. Moghaddasi, A. Fouman, M. E. Salehi, and M. Kargahi, ''Instruction-level NBTI stress estimation and its application in runtime aging prediction for embedded processors,'' *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 38, no. 8, pp. 1427–1437, Aug. 2019, doi: 10.1109/TCAD.2018.2846629.

[20] Y.-G. Chen, I.-C. Lin, and Y.-C. Wei, ''A novel NBTI-aware chip remaining lifetime prediction framework using machine learning,'' in *Proc. 22nd Int. Symp. Qual. Electron. Design (ISQED)*, Apr. 2021, pp. 476–481, doi: 10.1109/ISQED51717.2021.9424356.

[21] T. Siddiqua, S. Gurumurthi, and M. R. Stan, ''Modeling and analyzing NBTI in the presence of process variation,'' in *Proc. 12th Int. Symp. Qual. Electron. Design*, Mar. 2011, pp. 1–8.

[22] S. Han and J. Kim, ''NBTI-aware statistical timing analysis framework,'' in *Proc. 23rd IEEE Int. SOC Conf.*, Sep. 2010, pp. 158–163.

[23] S. Khan and S. Hamdioui, ''Modeling and mitigating NBTI in nanoscale circuits,'' in *Proc. IEEE 17th Int. On-Line Test. Symp.*, Jul. 2011, pp. 1–6.

[24] F. Firouzi, S. Kiamehr, and M. B. Tahoori, ''Statistical analysis of BTI in the presence of process-induced voltage and temperature variations,'' in *Proc. 18th Asia South Pacific Design Autom. Conf. (ASP-DAC)*, Jan. 2013, pp. 594–600.

[25] B. Aiguo and J. Li, ''A learning-based framework for circuit path level NBTI degradation prediction,'' *Electronics*, vol. 9, no. 11, pp. 1976–1988, 2020. [Online]. Available: https://doi.org/10.3390/electronics9111976

[26] S. M. Ebrahimipour, B. Ghavami, H. Mousavi, M. Raji, Z. Fang, and L. Shannon, ''Aadam: A fast, accurate, and versatile aging-aware cell library delay model using feed-forward neural network,'' in *Proc. 39th Int. Conf. Comput.-Aided Design*, Nov. 2020, pp. 1–9.

[27] S. Pendyala, S. A. Islam, and S. Katkoori, ''Gate level NBTI and leakage co-optimization in combinational circuits with input vector cycling,'' *IEEE Trans. Emerg. Topics Comput.*, vol. 8, no. 3, pp. 738–749, Jul. 2020, doi: 10.1109/TETC.2018.2799739.

[28] N. Hellwege, N. Heidmann, M. Erstling, D. Peters-Drolshagen, and S. Paul, ''An aging-aware transistor sizing tool regarding BTI and HCD degradation modes,'' in *Proc. 22nd Int. Conf. Mixed Design Integr. Circuits Syst. (MIXDES)*, Jun. 2015, pp. 272–277.

[29] A. Gomez and V. Champac, ''A new sizing approach for lifetime improvement of nanoscale digital circuits due to BTI aging,'' in *Proc. IFIP/IEEE Int. Conf. Very Large Scale Integr. (VLSI-SoC)*, Oct. 2015, pp. 297–302.

[30] Y. Yu, J. Liang, Z. Yang, and X. Peng, ''NBTI and power reduction using a workload-aware supply voltage assignment approach,'' *J. Electron. Test.*, vol. 34, no. 1, pp. 27–41, Feb. 2018, doi: 10.1007/s10836-018-5707-z.

[31] A. P. Shah, D. Rossi, V. Sharma, S. K. Vishvakarma, and M. Waltl, ''Soft error hardening enhancement analysis of NBTI tolerant schmitt trigger circuit,'' *Microelectron. Rel.*, vol. 107, Apr. 2020, Art. no. 113617.

[32] F. Andres Gomez and V. Champac, ''Selection of critical paths for reliable frequency scaling under BTI-aging considering workload uncertainty and process variations effects,'' *ACM Trans. Design Autom. Electron. Syst.*, vol. 23, no. 3, pp. 1–22, Apr. 2018, doi: 10.1145/3177864.

[33] J. P. Fishburn and A. E. Dunlop, ''TILOS: A posynomial programming approach to transistor sizing,'' in *Proc. Best ICCAD*, 1985, pp. 326–328.

[34] D. Blaauw, K. Chopra, A. Srivastava, and L. Scheffer, ''Statistical timing analysis: From basic principles to state of the art,'' *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 4, pp. 589–607, Apr. 2008.

[35] K. B. Sutaria, J. B. Velamala, A. Ramkumar, and Y. Cao, ''Compact modeling of BTI for circuit reliability analysis,'' in *Circuit Design for Reliability*. New York, NY, USA: Springer-Verlag, 2015, pp. 93–119.

[36] A. E. Islam, N. Goel, S. Mahapatra, and M. A. Alam, ''Reaction-diffusion model,'' in *Fundamentals of Bias Temperature Instability in MOS Transistors*. India: Springer, 2016, pp. 181–207.

[37] S. Jin, Y. Han, H. Li, and X. Li, ''$P^{(2)}$ CLRAF: An pre-and post-silicon cooperated circuit lifetime reliability analysis framework,'' in *Proc. 19th IEEE Asian Test Symp.*, Dec. 2010, pp. 117–120.

[38] S. Bian, M. Shintani, S. Morita, H. Awano, M. Hiromoto, and T. Sato, ''Workload-aware worst path analysis of processor-scale NBTI degradation,'' in *Proc. 26th Ed. Great Lakes Symp. (VLSI)*, May 2016, pp. 203–208.

[39] S. Duan, B. Halak, and M. Zwolinski, ''An ageing-aware digital synthesis approach,'' in *Proc. 14th Int. Conf. Synth., Modeling, Anal. Simulation Methods Appl. Circuit Design (SMACD)*, Jun. 2017, pp. 1–4.

[40] V. G. Rao and H. Mahmoodi, ''Analysis of reliability of flip-flops under transistor aging effects in nano-scale CMOS technology,'' in *Proc. IEEE 29th Int. Conf. Comput. Design (ICCD)*, Oct. 2011, pp. 439–440.

[41] C. Cerrone, R. Cerulli, and B. Golden, ''Carousel greedy: A generalized greedy algorithm with applications in optimization,'' *Comput. Oper. Res.*, vol. 85, pp. 97–112, Sep. 2017.

[42] A. Assad and K. Deep, ''A hybrid harmony search and simulated annealing algorithm for continuous optimization,'' *Inf. Sci.*, vol. 450, pp. 246–266, Jun. 2018.

[43] E. Maricau and G. Gielen, *Analog IC Reliability in Nanometer CMOS*. New York, NY, USA: Springer, 2013.

[44] National Interactive Maths Olympiad (NIMO) Group. (2005). *Predictive Technology Model (PTM)*. Accessed: Jan. 16, 2020. [Online]. Available: http://ptm.asu.edu/

[45] S. Ercolani, M. Favalli, M. Damiani, P. Olivo, and B. Ricco, ''Estimate of signal probability in combinational logic networks,'' in *Proc. 1st Eur. Test Conf.*, Jan. 1989, pp. 132–138.

[46] A. Jafari, M. Raji, and B. Ghavami, ''Impacts of process variations and aging on lifetime reliability of flip-flops: A comparative analysis,'' *IEEE Trans. Device Mater. Rel.*, vol. 19, no. 3, pp. 551–562, Sep. 2019.

MOHSEN RAJI received the Ph.D. degree in computer engineering from Amirkabir University of Technology, Tehran, Iran, in 2015. He is currently an Associated Professor with the School of Electrical and Computer Engineering, Shiraz University. His current research interests include dependable computing, reliable and robust logic designs, design automation of digital systems, and embedded systems.

REZA MAHMOUDI received the M.Sc. degree in computer architecture engineering from Shiraz University, Shiraz, Iran, in 2018. He is currently a Researcher at Shiraz University. His research interests include fault tolerant systems and electronic design automation.

BEHNAM GHAVAMI (Member, IEEE) received the Ph.D. degree in computer engineering from Amirkabir University of Technology, Tehran, Iran, in 2010. He is currently an Associate Professor with Shahid Bahonar University of Kerman (SBUK). He has authored and coauthored more than 50 technical articles in reputed journals and conference proceedings. His current research interests include design automation of digital systems, reliable and robust logic design, computer architecture, statistical analysis, and asynchronous logics.

SAEED KESHAVARZI received the M.Sc. degree in computer architecture engineering from Shiraz University, Shiraz, Iran, in 2017. He is currently a Researcher at Shiraz University. His research interests include fault tolerant systems and electronic design automation.