

TDTB Error Detecting Latches: Timing Violation Sensitivity Analysis and Optimization

Matheus T. Moreira^{1,2}, Dylan Hand², Peter A. Beerel^{2*}, Ney L. V. Calazans¹

¹Pontificia Universidade Católica do Rio Grande do Sul – Porto Alegre, Brazil

²University of Southern California – Los Angeles, United States

{matheus.moreira, ney.calazans}@puccrs.br, {dhand, pabeerel}@usc.edu

Abstract

Increasing process variations and sensitivity to operating conditions are making the design of traditional synchronous circuits a challenging task. Correct operation of these circuits relies on timing margins, which have an undesirably high cost in performance and power. One approach to mitigate this cost that is gaining substantial interest is the use of timing resilient microarchitectures that utilize error detecting sequential circuits. We evaluate the sensitivity of the transition detector with time borrowing error detecting latch to timing violations, including violations caused by glitches. Results show that the classic design is more constrained than previously believed and does not guarantee safe operation, i.e. does not guarantee that all timing violations will be captured. To overcome this limitation, we propose transistor level optimizations that enable safe operation, guaranteeing that all timing violations are captured, for a cost of 3 extra transistors, 30% in leakage power and 8% in energy.

Keywords

Error detecting latches, timing violation, glitch sensitivity

1. Introduction and Motivation

As silicon technologies scale into ultra-deep submicron nodes, process, voltage and temperature (PVT) variations play a crucial role in integrated circuits (ICs) design [1]. Process variations can occur across regions of the same die, from die to die and from wafer to wafer. Even if two transistors had precisely the same characteristics after fabrication, these characteristics would diverge over their life span because they may not have the same switching activity and will suffer differently from effects like aging [2]. Moreover, variations on the operating conditions of these devices result in substantial variation in their electrical characteristics [3]. In other words, delay and power variations are inevitable and increasingly problematic in semiconductors. To cope with such phenomena, contemporary clocked designs require delay margins. However, this compromises operating frequency, due to margins required to meet timing constraints under PVT, and energy efficiency, due to higher voltages being used to ensure timing closure [4].

Various approaches have been proposed to date for alleviating these problems, such as the addition of on-die voltage and temperature sensors and adaptive circuit techniques, but these solutions still rely on additional timing

guard bands [5]. A more promising approach to mitigating the guard bands is to use error detection sequential circuits (EDSs). These circuits allow resilient architectures [5] to operate at frequencies higher than those restricted by combinational path worst-case delays, by monitoring timing faults, also called errors. When timing faults are detected, they must be corrected, incurring extra clock cycle penalties. Subsequently, there is a tradeoff between the gains in performance and the increases in error rate when defining the operating frequency of circuits based on EDSs. This technique is gaining attention and various works report different approaches for designing EDSs and their benefits, including improving performance and increasing yield [4]-[13].

Of the EDS circuits published thus far, the transition detector with time borrowing (TDTB) error detecting latch (EDL), proposed by Bowman et al. in [5] is one we consider to be particularly interesting. For instance, these authors showed that the TDTB provides the best energy efficiency among several related circuit options. More importantly, the TDTB stands out by easing the task of dealing with metastability, which can be tricky in contemporary technologies, by preventing possible metastable signals from propagating through the datapath. Instead, these signals propagate to the control block, where they can be more easily handled. As far as we could verify, other than the seminal work presented in [5], only two works, [8] and [9], are available in literature that address the usage and design of TDTBs. However, none of these evaluated the sensitivity of the circuit to glitches, which we believe can jeopardize circuit functionality if not detected and signaled as errors.

In this article, we explore hazardous scenarios for the TDTB, employ an analytical model to explore its behavior for different timing violations, including glitches, and quantify its sensitivity to such effects. Accordingly, we discuss how designing a TDTB that ensures safe operation is not a straightforward task and propose optimizations, providing a guideline for TDTB designers. These optimizations rely on both transistor sizing and classic asynchronous design techniques that have long been employed by quasi-delay-insensitive designers. The optimized versions allow more relaxed design and ensure that all timing violations are captured. The cost for that is a 30% increase in leakage, 8% in energy per operation and 3 extra transistors.

2. TDTB State-of-the-Art

Bowman et al. proposed a new EDS called TDTB [5], which schematic is shown in Figure 1(a). They compare the TDTB with other EDSs through a test-chip in a 65nm

* Peter A. Beerel is also Chief Scientist, Technology Development at Intel, Calabasas, CA 91302.

CMOS technology and report that this is the most power efficient design. Also, they report that another advantage of the TDTB is that it enables removing metastability from the data path, moving it to the control path (more precisely in the error signal E) where it can be more easily treated. This characteristic makes the TDTB more interesting than other EDSs for contemporary applications, as metastability plays an increasingly important role in IC design. The circuit is composed of a latch, a transition detector (TD) (1) and an error latch (EL) (2), equivalent to the asymmetric C-Element shown in Figure 2 [14].

Figure 1(b) shows a timing diagram of the TDTB under normal operation. When Clk is high, the latch is transparent and the logic value of D is copied to Q, as shown in transitions D0 and D1. Whenever input D switches, the XOR gate generates a pulse on X due to the delay between its two inputs; see transitions E0 and E1. This delay is created by a delay element between D and dl and is referred throughout this text as Δ . If the pulse in X occurs while the latch is opaque, it does not affect the error signal, E, as in transition E0. Therefore, D must be stable before Clk becomes 1 and must remain stable throughout the high phase of Clk. However, if a transition on D (and subsequently a pulse on X) occurs while the latch is transparent, this represents a timing violation that must be detected by the TD (1) and stored in EL (2); see transition E1. Accordingly, the pulse generated on X is stored in the EL throughout the high phase of Clk, due to the memory scheme created by transistors MN3-4 and MP2-3. This error must be treated by the architecture before the latch becomes opaque, as E will return to 0 on the falling edge of Clk; see transition E2.

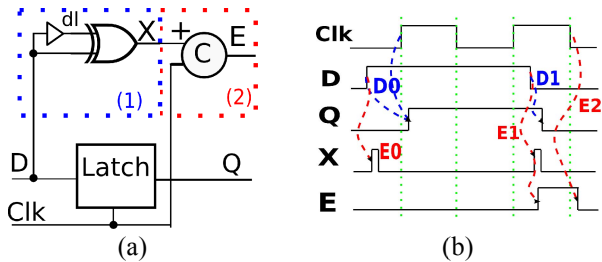


Figure 1: TDTB (a) schematic; (b) waveform [5].

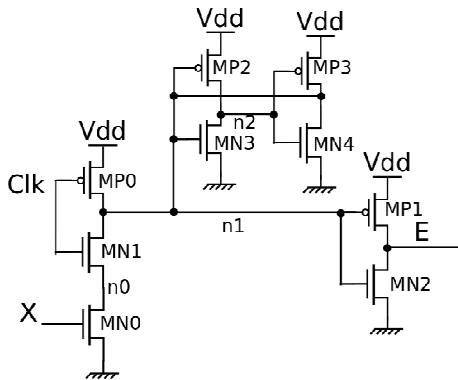


Figure 2: Semi-static asymmetric C-element schematic.

After its initial proposal, the TDTB was used for sub-threshold operation and error detection, correction and prediction techniques. In [8], the authors discuss optimizations at the architectural level using the TDTB for speculating the occurrence of timing errors. However, there is no analysis or

optimization of the TDTB circuit itself. Another work of particular relevance appears in [9]. There, Turnquist et al. propose a modification in the TDTB for subthreshold operation. The authors begin their discussion raising important concerns about the design of TDTBs. They show that dimensioning the transistors of the EL is a complex task, as the inverters loop that keeps node n1 stable must be weak enough so that pulses in X can switch its state. Furthermore, the correct behavior of this mechanism is easily corruptible by effects such as PVT variations and crosstalk, which makes it inadequate for subthreshold design. To overcome this problem the authors propose adding a transmission gate for opening the feedback loop when the clock is high, i.e. when the circuit is monitoring timing violations. The drawback is that, by using such a mechanism, the operation of the TDTB relies on internal capacitances and behaves as a dynamic gate. This characteristic requires extra care in its design and definition of constraints, and can be sensitive to PVT variations and other electrical phenomena [15].

3. Problem Statement

Recalling Figure 1(a) and (b), timing violations in D generate a pulse in X, and this pulse must be wide enough to signal an error in E. The width of this pulse can be easily adjusted by tuning Δ , where the bigger this delay is the bigger the pulse in X will be. However, this analysis is only valid for the cases where transitions in D are at least a time Δ apart. Scenarios where the timing violation is a glitch in D that is faster than Δ have not been previously evaluated, and we believe they are a potential source of timing failures. The correct operation of TDTB-based circuits relies on the premise that the error signal guarantees late arriving data does not exceed the max_delay constraints, which are defined to meet the latch setup constraints. However, if a glitch in D propagates through the TDTB latch but is not detected and signaled by E, it can exceed the specified timing constraints, jeopardizing the functionality of the circuit and allowing undesirable metastability to be injected in the datapath.

Assuming that a glitch with width DG occurs in D while the TDTB is monitoring errors, we identify three possible scenarios: (i) if $DG > \Delta$, two pulses with width Δ separated by $DG - \Delta$ are generated in X; (ii) if $DG = \Delta$, one pulse with width 2Δ is generated in X; (iii) if $DG < \Delta$, two pulses with width DG separated by $\Delta - DG$ are generated in X. Considering scenarios (i) and (ii), if Δ is defined as the minimum pulse that the EL can sense, an error signal will always be generated for glitches under such conditions. This is because the propagated pulses for (i) and (ii) are going to be Δ and 2Δ , respectively. In this case, one can tune the EL to be sufficiently sensitive because Δ is a known value. However, to guarantee that an error is always recorded for scenario (iii), the EL must reliably sense pulses of width DG . This is a more challenging task because these pulses can be generated by different sources and can be narrower than 1 gate delay as discussed in detail in [16]. Albeit narrow, these pulses are still hazardous and must be captured by the EDL, as they can propagate through the datapath latch as timing violations. Hence, in a robust design, the TD+EL, the block composed by (1) and (2), must be at least as sensitive to glitches as the latch. This

covers scenarios (i)-(iii) and guarantees that a glitch that propagates from D to Q will generate a pulse in X that the EL is able to detect and switch E to 1.

To verify the robustness of the TDTB we adapted the work proposed by Gili et al. in [17], an analytical approach of modeling the glitch sensitivity of combinational gates. The model relies on fitting simulation data obtained through simulation of the circuits under evaluation (CUEs) to a three-dimensional surface. Additionally, the authors present V_0 as the pulse height at the input of a CUE for a specific pulse width such that V_{out} , the height of the pulse generated at the output of the CUE, is equal to $V_{dd}/2$. In other words V_0 represents the switching threshold of CUEs, i.e. the minimum height for an input pulse width that creates a pulse in the output. In [17], the authors define V_0 as:

$$V_0 = V_{DC} \left(1 + \left(\frac{t_d}{t_{win}} \right)^\alpha \right) \quad (1)$$

where α is a curve-fit parameter and t_{win} is the input pulse width. t_d and V_{DC} are the propagation delay and switching threshold, respectively, and come from simulation.

Because we are primarily concerned with quantifying the switching threshold of CUEs, we focus on V_0 . In fact, for our CUEs, V_0 quantifies the minimum pulse width and height combination that causes the propagation of a timing violation to the datapath for the latch and that enables flagging an error in E for the TD+EL. The drawback is that V_{DC} cannot be easily determined with a high level of accuracy for the TD+EL because a DC analysis does not capture the sequential behavior of the TD created by Δ . Therefore, we introduce β to replace V_{DC} and allow the curve-fitting algorithm to determine this value based on our simulation data. Accordingly, we define V_0' as:

$$V_0' = \beta \left(1 + \left(\frac{t_d}{t_{win}} \right)^\alpha \right) \quad (2)$$

To collect data, we designed a TDTB targeting a 65nm bulk CMOS technology using conventional cells from the core library and a C-element from the ASCEnD library [18]. We then analyzed the sensitivity of the latch and the TD+EL to timing violations on D. For the latter we implemented two versions, with Δ delays of 4 and 6 inverters (4I and 6I). Note that preliminary simulation showed that smaller Δ delays lead to a pulse in X that is not wide enough to be captured by the EL. Figure 3 shows the simulation environment, defined according to the guidelines of [17], where each CUE receives input from an inverter and drives an output load. In our case we used a load equivalent to a fanout of 4. An ideal voltage source V feeds the input inverter and a capacitance C was placed between the inverter and the CUE to control the pulse height and width.

To collect data for curve fitting, we varied the width and height of pulses in V from 1ps to 200ps and 50mV to 1V (nominal voltage), respectively, and C from 1fF to 15fF. The combination of values enabled the analysis of over 90,000 glitch scenarios. This allowed a comprehensive exploration of the behavior for the evaluated circuits, enhancing the precision of curve fitting. We simulated all scenarios for each circuit using Cadence Spectre and measured the height and width of the glitch generated by the input inverter,

together with the pulse propagated through the CUE. Curve fitting was completed using Matlab's `lscurvefit` function, although any general curve fitting method should provide similar results. Through simulation and fitting (2), we obtained Figure 4 with the following parameters: for the latch, $\alpha=0.8029$ and $\beta=0.4409$ with $t_d=47.4ps$ and $R^2=0.9914$; for the 4I TD+EL, $\alpha=0.4963$ and $\beta=0.3196$ with $t_d=111.3ps$ and $R^2=0.9933$; and for the 6I TD+EL, $\alpha=0.7024$ and $\beta=0.3284$ with $t_d=122.8ps$ and $R^2=0.9752$.

Observing Figure 4, the most sensitive design will have a curve as close to the (0, 0) intersection as possible, indicating that it responds to input glitches that present small heights and widths. Additionally, this analysis is not only useful for timing violations caused by glitches, as full transitions will also generate a pulse in X and the sensitivity of the EL will determine its capability of capturing the violation. As Figure 4 shows, the 6I TD+EL is clearly not suited for safe operation, as its sensitivity is worse than that of the latch. The 4I TD+EL, on the other hand, safely captures small glitches, but presents similar sensitivity to wide glitches as the latch. As discussed in [16], glitches propagated through combinational logic can have different widths and heights. Therefore, the obtained results indicate that the classic TDTB does not guarantee safe operation, i.e. some glitches could be propagated through the latch without generating an error signal.

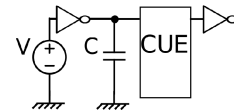


Figure 3: Simulation environment setup from [17].

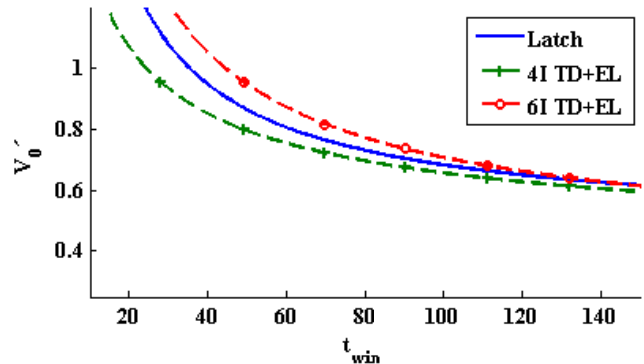


Figure 4: V_0' for TDTB.

4. TDTB Optimization

In view of the problem described in Section III, we propose two optimizations for the TDTB: (i) transistor level optimization of the XOR in the TD for enlarging the pulse generated in X; and (ii) the use of a static C-Element implementation for ELs. Throughout this text we refer to (i) as the optimized XOR TD (OX-TD) and (ii) as the static EL (S-EL). Recalling Figure 1(a), the XOR gate generates pulses that feed the EL whenever D switches. The wider these pulses are the easier it is for them to be captured. In this way, it is desirable that the XOR gate present: (1) a fast response to transitions in D; (2) a slow response to transitions in d_l ; (3) fast low-to-high output transition arcs; and (4) slow high-to-low output transition arcs. Figure 5 shows the schematic of the XOR circuit we employed in our

experiments. The implementation is a classic complementary design that consists of a NOR2 connected to an AOI22 gate. We chose this topology instead of a pass-transistor logic-based gate to avoid charge sharing effects, which could compromise glitch sensitivity. While this particular design was available in our library, the strategy described herein can be applied to other topologies as well.

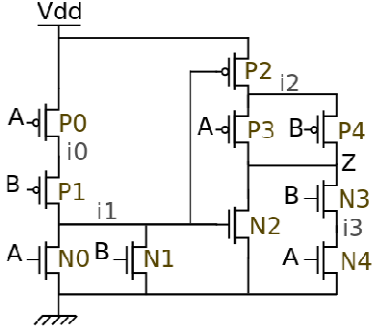


Figure 5: XOR circuit schematic.

In our design, D was connected to input B of the XOR (see Figure 5) because this is the most responsive input, which is in agreement with item (1) of the TD optimization. Furthermore, we modified the width of transistors P2 and P4 to make the gate even more responsive to transitions in D. Note that a tradeoff exists in setting the transistors' widths, as a wider gate increases the driving strength but also increases its input capacitance, which makes it less responsive to glitches. Accordingly, we employed SPICE to sweep the widths and defined these as the largest size before the input capacitance became dominant. The same approach was applied to transistor N1. These optimization steps enabled meeting items (1) and (3) of the TD optimization, as the XOR became more responsive to transitions on D and low-to-high output transition arcs were sped up due to the increased width. Note that another possible optimization vector is re-placing P2 and P4 with low V_t transistors. While this does increase sensitivity to D, it also causes an increase in leakage power, making it undesirable for low power designs. Next, we reduced the width of transistors N2-4 to minimum size while also reducing their responsiveness to transitions on D1 and increasing the delay of high-to-low output transition arcs. This facilitates satisfying items (2) and (4) of the TD optimization.

Using an OX-TD, we designed an optimized TD+EL, referred to as OX-TD+EL. The same analytical model described in Section III was employed to obtain parameters $\alpha=0.3914$, and $\beta=0.2510$ with $t_d=95.9\text{ps}$ and $R_2=0.976$. With these values, we obtained a closed form solution for V_0' as a function of t_{win} . As Figure 6 shows, the optimized circuit is more sensitive to glitches than the latch. In fact, it is able to capture all glitches propagated by the latch. Taking V_0' measured at t_{win} of 50ps and 100ps as a metric for sensitivity, this circuit is respectively 33% and 27% more sensitive than the original TD+EL. However, this improvement comes at a cost in average leakage power and energy per operation. The original TDTB achieves $0.181\mu\text{W}$ and 18.51fJ , respectively, while the OX-TDTB reaches $0.229\mu\text{W}$ and 21.6fJ , which gives respective overheads of 27% and 17%.

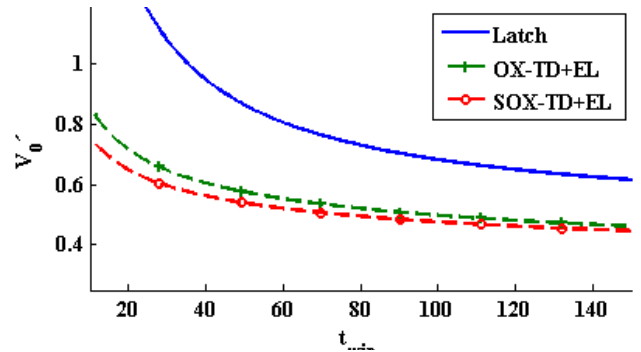


Figure 6: V_0' for latch, OX-TD+EL and SOX-TD+EL.

Further sensitivity optimizations and overhead reductions can be obtained using an S-EL, optimization (ii). For this optimization, we analyzed the C-element used in the original TDTB, which employs a semi-static topology in which a conflict between the keeper and forward path exists while the gate is switching [18]. Analyzing the schematic in Figure 2, for switching E to 0, the PMOS of the logic stack (MP0) must overpower the NMOS of the feedback inverter (MN4), as both drive the internal node n1. Similarly, to switch E to 1, transistors MN0 and MN1 must overpower the PMOS of the feedback inverter (MP3). Such characteristics rely on careful design of the C-element, since transistors MP0, MN0 and MN1 must always be stronger than transistors MP3 and MN4. A static implementation, which schematic appears in Figure 7, alleviates these problems. The basic difference from the original circuit is the addition of transistors MN5, MP4 and MP5. These avoid the conflict that is present in the semi-static topology, as they disconnect the feedback inverter from the power rails while the output is allowed to switch.

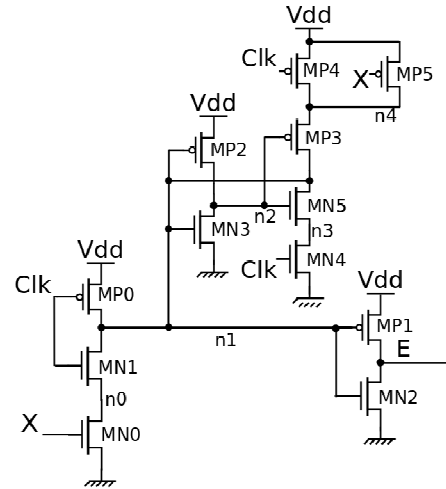


Figure 7: Static C-element schematic.

For instance, consider that output E is at 1 and the input Clk switches from 1 to 0. In this case, MP0 is turned on and starts to charge the internal node n1. At the same time, MN5 is turned off, avoiding the feedback inverter to discharge n1, preventing the conflict situation. Also, as soon as n1 is charged, n2 is discharged through MN3, storing the value in the memory mechanism. Now consider that X is stable at 0 and Clk switches back to 1. In this case, we have the value of n1 being kept by the path composed by MP3 and MP5, as n2 and X are at 0. However, as soon as a glitch is detected

and X switches to 1, n1 is discharged through MN0 and MN1, switching the error signal E to 1. Note that in this case conflict is also avoided, because as soon as X switches to 1, MP5 is turned off, disconnecting the feedback path. This technique is common in asynchronous designs and is well known to provide improvements in operating speed, leakage power and energy [18].

Using an S-EL, we designed an optimized OX-TD+EL designated as the SOX-TD+EL. For this circuit, the glitch sensitivity model generated $\alpha=0.3499$ and $\beta=0.2485$ with $t_d=77.9\text{ps}$ and $R2=0.9982$. V_0' as a function of t_{win} for this circuit is also plotted in Figure 6. As the chart shows, it provides even better sensitivity than the OX-TD-EL, 37.6% and 30.3% higher than the latch at t_{win} of 50ps and 100ps, respectively. Moreover, it enables modest reductions on power and energy overheads. Accordingly a SOX-TDTB presents a leakage power of $0.236\mu\text{W}$ (an overhead of 30% over the original design) and 20fJ for energy per operation (an overhead of 8% over the original design), at the cost of 3 extra transistors.

5. Conclusions

This work addressed the sensitivity of the TDTB EDL to timing violations, including those caused by glitches. Our analysis shows that the classic implementation is not sufficient for ensuring safe operation. It allows the propagation of undetected errors in the datapath, possibly leading to metastable states. In order to overcome this problem, we proposed two optimizations: OX-TDTB and SOX-TDTB. The optimized circuits are able to detect all violations that could be propagated by the latch into the datapath, ensuring safe operation, at the cost of 30% increase in leakage, 8% in energy per operation and 3 extra transistors

6. Acknowledgements

This work was partially supported by CNPq (under grants 55679/2009-6, 310864/2011-9, 401839/2013-3 and 200147/2014-5) and by CAPES under grant 2129/14-0.

7. References

- [1] K. J. Kuhn, M. D. Giles, D. Becher, P. Kolar, A. Kornfeld, R. Kotlyar, S. T. Ma, A. Maheshwari, S. Mudanai, "Process Technology Variation," *Electron Devices, IEEE Transactions on*, 58(8), 2011, pp. 2197-2208.
- [2] E. Maricau and G. Gielen, "Transistor Aging-Induced Degradation of Analog Circuits: Impact Analysis and Design Guidelines," in *European Solid-State Circuits Conference*, 2011, pp. 243-246.
- [3] A. Rahimi, L. Benini, R. K. Gupta, "Analysis of Instruction-Level Vulnerability to Dynamic Voltage and Temperature Variations," in *Design, Automation and Test in Europe Conference and Exhibition*, 2012, pp. 1102-1105.
- [4] S. Kim; I. Kwon; D. Fick, K. Myungbo C. Yen-Po D. Sylvester, "Razor-lite: A side-channel error-detection register for timing-margin recovery in 45nm SOI CMOS," in: *IEEE International Solid-State Circuits Conference*, 2013, pp. 264-266.
- [5] K. A. Bowman, J. W. Tschanz, K. N. Sung, J. C. Lee, C. B. Wilkerson, S. L. Lu, T. Karnik, V. K. De, "Energy-Efficient and Metastability-Immune Resilient Circuits for Dynamic Variation Tolerance," *IEEE Journal of Solid-State Circuits*, 44(1), 2009, pp. 49-63.
- [6] M. Fojtik, D. Fick, Y. Kim, N. Pinckney, D. Harris, D. Blaauw, D. Sylvester, "Bubble razor: An architecture-independent approach to timing-error detection and correction," in *IEEE International Solid-State Circuits Conference*, 2012, pp. 488-490.
- [7] Z. Guowei and P. A. Beerel, "Stochastic analysis of Bubble Razor," in *Design, Automation and Test in Europe Conference and Exhibition*, 2014, pp. 24-28.
- [8] Y. Shi, H. Igarashi, N. Togawa, M. Yanagisawa, "Suspicious Timing Error Prediction with in-cycle Clock Gating," in *International Symposium on Quality Electronic Design*, 2013, pp. 335-340.
- [9] M. J. Turnquist and L. Koskinen, "Sub-threshold Operation of a Timing Error Detection Latch," in *PhD Research in Microelectronics and Electronics*, 2009, pp. 124-127.
- [10] K. A. Bowman, J. W. Tschanz, S. L. Lu, P. A. Aseron, M. M. Khellah, A. Raychowdhury, B. M. Geuskens, C. Tokunaga, C. B. Wilkerson, T. Karnik, V. K. De, "A 45 nm Resilient Microprocessor Core for Dynamic Variation Tolerance," *IEEE Journal of Solid-State Circuits*, 46(1), Jan 2011, pp. 194-208.
- [11] T. Sato and Y. Kunitake, "A Simple Flip-Flop Circuit for Typical-Case Designs for DFM," in *International Symposium on Quality Electronic Design*, 2007, pp. 539-544.
- [12] Y. Qiaoyan and D. Stock, "Collaborative Error Control Method for Sequential Logic Circuits," in *International Symposium on Circuits and Systems*, 2013, pp. 785-788.
- [13] T. Sato and Y. Kunitake, "Critical Issues Regarding A Variation Resilient Flip-Flop," in *Workshop on Synthesis and System Integration of Mixed Information Technologies*, 2007.
- [14] M. T. Moreira, B. Oliveira, J. Pontes, F. Moraes, N. Calazans, "Adapting a C-element design flow for low power," in *International Conference on Electronics, Circuits and Systems*, 2011, pp. 45-48.
- [15] J. M. Rabaey, A. Chandrakasan, B. Nikolic, "Digital Integrated Circuits: A Design Perspective," Prentice Hall, 2003, 761 p.
- [16] R. Garg, C. Nagpal, S. P. Khatri "A fast, analytical estimator for the SEU-induced pulse width in combinational designs," In: *Annual Conference on Design Automation*, 2008, pp. 908-918.
- [17] X. Gili, S. Barcelo, S. A. Bora, J. Segura "Analytical Modeling of Single Event Transients Propagation in Combinational Logic Gates," *IEEE Transactions on Nuclear Science*, 59(4), 2012, pp. 971-979.
- [18] M. T. Moreira, B. S. Oliveira, F. G. Moraes, N. L. V. Calazans, "Impact of C-elements in asynchronous circuits," in *International Symposium on Quality Electronic Design*, 2012, pp. 19-21.