SOFT-ERROR RESILIENCE FRAMEWORK FOR RELIABLE AND ENERGY-EFFICIENT CMOS LOGIC AND SPINTRONIC MEMORY ARCHITECTURES

by

FARIS S. ALGHAREB
B.S. University of Mosul, 2007
M.S. University of Mosul, 2009

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Electrical and Computer Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2019

Major Professor: Ronald F. DeMara

ABSTRACT

The revolution in chip manufacturing processes spanning five decades has proliferated high performance and energy-efficient nano-electronic devices across all aspects of daily life. In recent years, CMOS technology scaling has realized billions of transistors within large-scale VLSI chips to elevate performance. However, these advancements have also continually augmented the impact of Single-Event Transient (SET) and Single-Event Upset (SEU) occurrences which precipitate a range of Soft-Error (SE) dependability issues. Consequently, soft-error mitigation techniques have become essential to improve systems' reliability. Herein, first, we proposed optimized soft-error resilience designs to improve robustness of sub-micron computing systems. The proposed approaches were developed to deliver energy-efficiency and tolerate double/multiple errors simultaneously while incurring acceptable speed performance degradation compared to the prior work. Secondly, the impact of Process Variation (PV) at the Near-Threshold Voltage (NTV) region on redundancy-based SE-mitigation approaches for High-Performance Computing (HPC) systems was investigated to highlight the approach that can realize favorable attributes, such as reduced critical datapath delay variation and low speed degradation. Finally, recently, spin-based devices have been widely used to design Non-Volatile (NV) elements such as NV latches and flip-flops, which can be leveraged in normally-off computing architectures for Internet-of-Things (IoT) and energy-harvesting-powered applications. Thus, in the last portion of this dissertation, we design and evaluate for soft-error resilience NV-latching circuits that can achieve intriguing features, such as low energy consumption, high computing performance, and superior soft errors tolerance, i.e., concurrently able to tolerate Multiple Node Upset (MNU), to potentially become a mainstream solution for the aerospace and avionic nanoelectronics. Together, these objectives cooperate to increase energy-efficiency and soft errors mitigation resiliency of larger-scale emerging NV latching circuits within iso-energy constraints. In summary, addressing these reliability concerns is

paramount to successful deployment of future reliable and energy-efficient CMOS logic and spintronic memory architectures with deeply-scaled devices operating at low-voltages.

Dedicated to my lovely family

ACKNOWLEDGMENTS

I would like to thank my advisor, Prof. Ronald F. DeMara, for his strong support, guidance, and encouragement that truly guided me to achieve this dissertation. I really appreciate his passion and dedication to research and feel proud and fortunate to have worked with him throughout my graduate studies.

I extend my thanks to the Ph.D. dissertation committee members: Dr. Mingjie Lin, Dr. Sumit Jha, Dr. Cliff Zou, and Dr. Zixia Song for their time to evaluate the contents of this dissertation and provide valuable comments. I also would like to express my deep appreciation to Diana Camerino for her kind help throughout my course of study.

I would also like to acknowledge the financial support from the Ministry of Higher Education and Scientific Research in Iraq (MOHESR). Their support was fundamental for the completion of the Ph.D. program requirements; for that I am deeply grateful.

My sincere thanks are due to my family, relatives and friends for their continuous support, especially my mother, uncle Dr. Khamees, sister Dr. Nibras, and brother Dr. Safwan for their support, inspiration, and love that assisted making my dreams came true. Other brothers and sisters for their many years of support and love during my studies.

Finally, I would like to thank other students in CAL group. I really enjoyed working with all of you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AES Advanced Encryption Standard

ALU Arithmetic Logic Unit

ASIC Application-Specific Integrated Circuit

BISER Built-In Soft-Error Resilience

CLK Clock

CMFs Common Mode Failures

CMOS Complementary Metal Oxide Semiconductor

DICE Dual Interlocked-storage Cell

DMR Dual Modular Redundancy

DNU Double Node Upset

DSP Digital Signal Processing

DVFS Dynamic Voltage Frequency Scaling

ECC Error Correcting Code

EDP Energy Delay Product

FF Flip-Flop

FCER Fault Coverage Energy Ratio

FinFETs Fin-typed Field-Effect Transistors

FIT Failure In Time

FPGAs Field-Programmable Gate Arrays

HPC High-Performance Computing

HSTR-DEC Hybrid Spatial-Temporal Redundancy-Double Error Correction

LCM Local Clock Manager

LET Linear Energy Transfer

MBU Multi-Bit Upset

MTJ Magnetic Tunnel Junction

NMR N-Module Redundancy

NTV Near-Threshold Voltage

NV Non-Volatile

PCSA Pre-Charge Sense Amplifier

PIPB Propagation-Induced Pulse Broadening

PTM Predictive Technology Model

PV Process Variation

RCC Reinforcing Charge Collection

RDF Random Dopant Fluctuation

RHBD Radiation-Hardening By Design

RHBP Radiation-Hardening By Process

SDC Silent Data Corruption

SE Soft-Error

SEE Soft Error Effect

SEMT Single-Event Multiple Transients

SER Soft Error Rate

SEFI Single-Event Functional Interrupt

SET Single-Event Transient

SEU Single-Event Upset

SFC Soft Fault Coverage

SHE Spin Hall-Effect

SOI Silicon-On-Insulator

STT Spin Transfer Torque

SV Self-Voting

TDDB Time Dependent Dielectric Breakdown

TG Transmission Gate

TID Total Ionizing Dose

TMR Triple Module Redundancy

TSVL Temporal Self-Voting Logic

CHAPTER 1: INTRODUCTION

This dissertation evaluates and presents the necessity for radiation-induced soft-error mitigation techniques of nano-electronic computing system architectures. As technology trends of Complementary Metal Oxide Semiconductor (CMOS) devices have improved the capability of contemporary processors. However, it has increased the susceptibility of VLSI circuits to transient faults. Thus, Soft Error Rate (SER) masking approaches are sought to increase reliability within area, speed, energy consumption, and fault masking resilience constraints. In this dissertation, energy-efficient and cost-effective redundancy-based soft error mitigation techniques to achieve resilience in VLSI circuits and systems are proposed. In particular, resilient approaches are introduced to deal with radiation-induced transient effects, i.e., Single-Event Transient (SET) and Single-Event Upset (SEU), by carrying out tradeoffs between energy consumption and fault resilience. In addition, the effects of process variation at Near-Threshold Voltage (NTV) region on the redundancy-based soft error mitigation approaches were investigated to leverage the technique that achieves an increased performance, satisfies the user specified objective metrics, within an iso-energy constraints. Whereas, the later part of this dissertation seeks to cover reliability of deeply-scaled hybrid CMOS/spintronic (MTJs) nonvolatile latching circuits, due to the radiation-induced soft errors in the CMOS-based logic portion that is utilized for read/write access operations. A study to completely and more efficiently address these transient effects to improve reliability will be conducted to make this emerging technology feasible. This chapter discusses the necessity and motivation for techniques presented in this dissertation, and concludes with a list of contributions to the state-of-the-art as a result of this work.

1.1 Categories of Single-Event

As Complementary Metal Oxide Semiconductor (CMOS) technology continues to scale towards fundamental physical limits, its immunity against charged particle strikes decreases significantly. This implies that the continued scaling of CMOS technology process reduces the quantity of critical charge, Q_{crit}, transistor's threshold voltage, a node capacitance, and the supply voltage (V_{DD}). This gives rise to increase the susceptibility towards radiation-induced soft errors caused by numerous energetic particles, such as protons and heavy ions in space and alpha particles from packaging materials and cosmic rays neutrons at sea level [13]. Thus, soft errors at the chip level has become a major reliability challenge for VLSI designs [14]. In order to achieve high-reliability for contemporary computing systems, memory protection techniques, i.e., Error Correcting Code (ECC), have been utilized. However, such techniques are inadequate to completely and efficiently protect the reliability of VLSI circuits and systems operating at NTV. In particular, soft error effects (SEE) in logic circuits, i.e., latches/flip-flops and combinational logic, have become significant contributors to increase Soft Error Rate (SER) at the system-level [15, 16]. Therefore, it has become essential to ensure the integrity of logic paths by utilizing efficient soft-error mitigation techniques.

At the chip-level, various types of soft error effects can occur in CMOS logic circuits due to incident energetic particles, including 1) particles that hit inside combinational logic and generate a transient pulse, the Single-Event Transient (SET) [17], which propagates through downstream logic, and eventually it might be captured by a storage element, i.e., latch/flip-flop, thereby causing an upset, 2) energetic particles that strike node(s) inside a storage element to cause a Single-Event Upset (SEU) that flips the bit state in a DRAM memory cell or a SRAM-based register, resulting in a soft error, and 3) energetic particle strikes on global signal lines, such as control signals or instruction lines, can result in Single-Event Functional Interrupt (SEFI) that causes a temporary malfunction, i.e., interruption of normal operation [18], as the wrong instruction might

be executed. The corrupted data can be rewritten with new legitimate data and can circumvent the effect of incorrect state. When the physical device is not destructed, any such temporary glitches are considered to constitute soft errors. Altogether, these events can cause catastrophic failures in mission-critical applications. Thus, efficient protection schemes need to be evaluated with higher priority as moving towards NTC for energy-efficient operation.

1.2 Soft Error Rate in Nanometer-scaled VLSI Devices

Practically speaking, soft errors occur when an energetic charged particle strikes a sensitive node of a circuit. This induces electron-hole pairs that cause charge/discharge at the struck transistor junction [19], which depends on the amount of charge collected at the node. The effect of such transient pulse depends on its duration, the state of the struck device (ON/OFF), and the type of struck node, i.e. drain side [20]. If the collected charge exceeds the critical amount of charge, Q_{crit}, then it produces some momentary charge and flips the voltage state of some circuit node(s), thereby causing a temporary glitch condition. The rate of these errors defined as Soft Error Rate (SER), which is a measurement metric that is used to evaluate the sensitivity of a digital circuit to transient and upset faults [17].

1.3 Why Are Soft Errors Important? Needs for Soft-Error Resilience Logic Circuits

The revolution in chip manufacturing processes spanning five decades has proliferated high performance and energy-efficient nano-electronic devices across all aspects of daily life. In recent years, CMOS technology scaling has realized billions of transistors within large-scale VLSI chips to elevate performance. However, these advancements have also continually augmented the impact of SET and SEU occurrences which precipitate a range of Soft-Error (SE) dependability

issues [14, 21]. In the earlier technologies, soft errors were a major concern to the reliability of Integrated Circuits (ICs) for space applications. Moving forward to advanced technology node, the susceptibility of sub-micron electronics to radiation-induced transient effects has been of increased interest, and this phenomenon dominates the overall SER of large-scale logic circuits [1, 22]. In addition, the trend toward using smaller, denser, and faster CMOS devices has influenced logic circuits to experience larger leakage currents, thereby negatively impacting the reliability of deeply-scaled circuits by degrading their performance and robustness [23]. This is because the heat that results from large leakage currents can induce thermal noise, which can lead to bit upset in a memory element [24, 25]. Thus, as technology has advanced and transistors have become smaller, soft errors have become a major contributor to reliability issues at sea level [20].

Taking all of these reliability-impacting factors into consideration, soft errors have been assessed to dominate a variety of failure mechanisms, whereby the SER might exceed the Failure In Time (FIT) metric [16,26,27]. As listed in Table 1.1, the SER of logic and memory per chip has increased 100-fold with technology scaling, especially for recent generations. In addition, due to the increasing integration density and growing complexity of ICs, the number of susceptible nodes to transients faults have increased. This leads to a higher SER that has recently been considered a serious concern for reliable memory and logic circuits [14, 28]. If such transient errors are not budgeted or compensated for, then the circuit can fail to perform correctly. Therefore, soft-error suppression becomes essential to improve the reliability of resilient systems operating in harsh environments like space, marine, military, and even in commercial applications [16, 20, 26, 29].

1.4 FinFET Structure Devices

The High-κ/Metal Gate (HK-MG) devices with vertical channels, i.e., the fin-typed Field-Effect Transistors (FinFETs) have been introduced to continue scaling and also show improved low-

leakage and high performance over planar MOSFET devices [30, 31]. The tri-gate FinFET devices have been adopted by Intel at the 22nm technology node and beyond since they handle the problems of Short-Channel Effect (SCE) and provide more robust gate controllability than planar MOSFET devices, while offering improved switching characteristics for potential scalability beyond 10nm half-pitch [30, 32, 33]. Additionally, multi-gate bulk FinFET devices improve variations and thus the system's overall performance has significantly improved. Moreover, FinFET devices have exhibited higher tolerance to soft error effects, while being less prone to particle strikes compared to planar structure CMOS devices, thereby they have assisted in increasing the robustness of large-scale logic circuits [16,34]. In contrast, even though scaling the transistor volume helps in reducing the susceptible area, there exists an increased SER. This is due to two major factors including: the critical quantity of charge, Q_{crit}, of a node capacitance has aggressively diminished, while the other factor is that the density of integration is doubling the number of transistors per unit area over time, leading to an increased SER as low energy particles potentially can deposit an adequate amount of charge, i.e., sufficient Linear Energy Transfer (LET), to cause soft errors.

Table 1.1: Trends of SER as impacted by technology scaling [7, 8]. However, vertical channel devices, i.e., FinFETs, can offer alternatives that achieve increased tolerance to SER for technologies beyond 22nm. For instance, Intel recently released its processors within 14nm technology node.

Design Rule (nm)	**SER (Logic+Memory)**	**Total SER (Data Center)**	**Chip Count**
180	1	–	–
130	15	–	–
90	25	10^5	$10^{4.4}$
65	50	$10^{5.5}$	$10^{4.8}$
45	60	$10^{6.25}$	$10^{5.3}$
32	75	10^7	$10^{5.7}$
22	90	$10^{7.5}$	$10^{6.15}$
16	10^2	$10^{8.8}$	$10^{6.75}$

1.5 Contributions of the Dissertation

The dissertation's inspiration is partly to give an insight into the soft-error mitigation techniques and their challenges related to the design of reliable systems, to assist in selecting the proper technique for satisfying specific SER constraints. The main contributions of this dissertation are highlighted below.

Energy-efficiency vs fault resilience: In this work, a reliable and cost-effective, area and energy efficient, redundancy-based approaches have been developed to mitigate soft errors at the circuit/module-level. The proposed SE-tolerant techniques are able to tolerate both transient and upset faults and provide complete SEU protection, detection, and masking. The presented schemes impose comparable speed performance compared to the previously presented self-voting DMR hybrid redundancy approach. In addition, the proposed approach incurs reduced area overhead as compared to the conventional TMR, aside from its ability to tolerate double SEUs simultaneously. The efficacy of these approaches is quantified using Fault Coverage Power Ratio metric. This work is presented in Chapter 3.

Impact of PV at NTV on redundancy-based soft error mitigation techniques: The energy and performance costs, which are crucial design considerations due to prevalent use of low-voltage operation in current computing systems, were deeper investigated. In particular, detailed energy and performance tradeoffs of redundancy-based approaches for soft error mitigation in 16-nm non-planar FinFET and 45-nm planar CMOS structures are provided. For process variation, a comparison between 45-nm planar CMOS and 16-nm FinFET CMOS structure was carried out regarding energy-efficiency. Furthermore, the effects of both delay variation and technology node scaling on redundancy-based soft error mitigation techniques at NTV for iso-energy constraints have been revealed to depict the quantitative results. Additional important insight of this work was to identify and quantify the increase in energy cost of spatial redundancy at contemporary and future tech-

nology nodes such as 16 nm. Monte-Carlo simulation results demonstrate that energy-efficiency benefits of scaled technology devices (16-nm) as compared to 45-nm node due to increase in performance variations. The highlights of work presented in Chapter 4 are given below:

1. *Assessing Resilience vs Area of Soft-Error Masking schemes:* determining empirically the costs of spatial and temporal redundancy under the impact of technology scaling.

2. *Evaluating the Energy and Delay Cost of Redundancy at NTV:* determining the delay of redundant systems at NTV within a given energy budget.

3. *Soft-Error Resilience Sensitivity to Process Variation:* identifying the increased impact of σV_{th} on alternative SER reduction strategies for 45-nm planar and tri-gate 16-nm bulk FinFET CMOS devices.

4. *Fault Coverage Energy Ration:* employing the FCER metric as an optimization parameter for design of resilient circuits, and its applicability to guide circuit synthesis algorithms while meeting performance goals.

5. *Design Diversity and Spatial Redundancy:* quantify the pros and cons of design diversity for redundancy-based soft error mitigation approaches.

Soft-error resilient hybrid CMOS/spintronic STT-MTJ latching circuit: Emerging spin-based devices are under intensive developments and investigations to improve their reliability. In this work, we design and evaluate soft-error resilience hybrid CMOS/spin-based latching circuit for nonvolatile high-performance applications. The proposed approach can achieve intriguing features, such as low energy consumption, high computing performance, superior soft error effects resilience, i.e., concurrently can tolerate MNU, to potentially become as a mainstream solution for the aerospace and avionic nanoelectronics.

Soft-error resilient NVFF circuits for eneregy-aware and reliability-aware architectural state storage: In this chapter, radiation hardening Non-Volitale Flip-Flop (NVFF) circuits are developed using spin-based devices that deliver reduced power dissipation and optimized CLK-to-Q delay, while achieving complete soft errors masking coverage for both SEUs and DNUs. It identifies the need for energy-aware and reliability-aware soft-error resilient latching circuits to improve soft errors masking robustness of emerging large-scale NVFF circuits while concurrently reducing its power dissipation to realize energy efficiency. Specifically, the contributions of chapter 6 are as follows:

1. Develop an NV-latch using Spin Hall Effect (SHE)-based Magnetic Tunnel Junctions (MTJs) that can tolerate multiple node upsets.
2. Design of low energy and high-performance SEU and DNU tolerant CMOS-based latches.
3. Develop four different NVFFs, which can be utilized in normally-off and NV computing architectures.

CHAPTER 2: LITERATURE REVIEW

In this chapter, a wide range of resilient SER mitigation techniques that exploit specific attributes of different levels of a design for energy-savings to attain reliable protection capabilities are discussed. These techniques have been adopted to mitigate soft errors at different levels of abstraction. To highlight the characteristics of each technique and demonstrate the accuracy of soft error mitigation schemes, the protection techniques are categorized based on the design level of the protection scheme. The taxonomy of SER masking techniques relies on the Single Event Transient (SET) pulse generation at the device-level, propagation at the circuit-level, and capturing at the circuit/module-level. Also developed is a concise taxonomy for the sources of induced soft errors and variation. A comparison among the predominant mitigation schemes in terms of area, energy consumption, fault coverage, and design complexity is carried out to leverage the robustness of schemes that achieve the highest performance over the stated issues. Likewise, a review of the impact of technology and voltage scaling trends on the SER and its masking techniques are discussed. The inspiration of this survey chapter has been to provide a compendium of design insights for soft error mitigation techniques to identify the most efficient hardening schemes, in order to serve as a guide for researchers and designers of reliable CMOS circuits and systems.

2.1 SER-Induced System Dependability Issues

In the nanometric technologies, integrated circuits are associated with the diminishing of device feature size, massive growth in integration density, and lower supply voltages. Thus, these desirable attributes have raised the vulnerability of the integrated circuits to soft errors as they impose risks for terrestrial and space applications. Next, system reliability at different abstraction layers and the sources of transient and permanent faults are discussed.

2.1.1 Reliability at Multiple Abstraction Levels

A system's reliability can be impacted by different effects. Herein, we concentrate on the effects of soft errors as a major reliability issue in highly-scaled devices and systems. Suppression of soft errors can be achieved at different abstraction levels. As seen in Figure 2.1, a fault, error, and failure can occur at three distinguished layers. As depicted by the cumulative arc in Figure 2.1, during the mission each component may transition from viable status to faulty status for highly-scaled devices. This transition may occur due to cumulative effects of deep submicron devices such as *Time Dependent Dielectric Breakdown (TDDB)* [35], which is caused by an electric field weakening of the gate oxide layer, the *Total Ionizing Dose (TID)* of cosmic radiation, *Electromigration* within interconnections, and other progressive degradations over the mission lifetime. Meanwhile, transient effects such as incident alpha particles, which ionize critical amounts of charge, ground bounce, and dynamic temperature variations may cause either long lasting or intermittent reversible transitions between viable and faulty status. A fault may affect the data that is being processed and then impact the final result and produce a failure. However, transient faults can be remedied at the *Resource Layer*, allowing the next layer's input to be error-free, or faults may lie dormant whereby the physical resource is defective, yet currently unused. Later in the computations, dormant faults become active when such components are utilized, affecting the result of the resource layer.

The *Behavioral Layer* shown in Figure 2.1 depicts the outcome of utilizing viable and faulty physical components. Viable components result in correct behavior during the interval of observation. An error that occurs but does not incur any impact to the result of the computation is termed a silent error. Silent errors, such as a flipped bit due to a faulty memory cell at an address that is not referenced by the application or an error in a *Triple Module Redundancy (TMR)* system that occurs in a single instance but the final output is error-free as long as the other two inputs are correct, remain isolated at the Behavioral Layer without propagating to the application [36, 37].

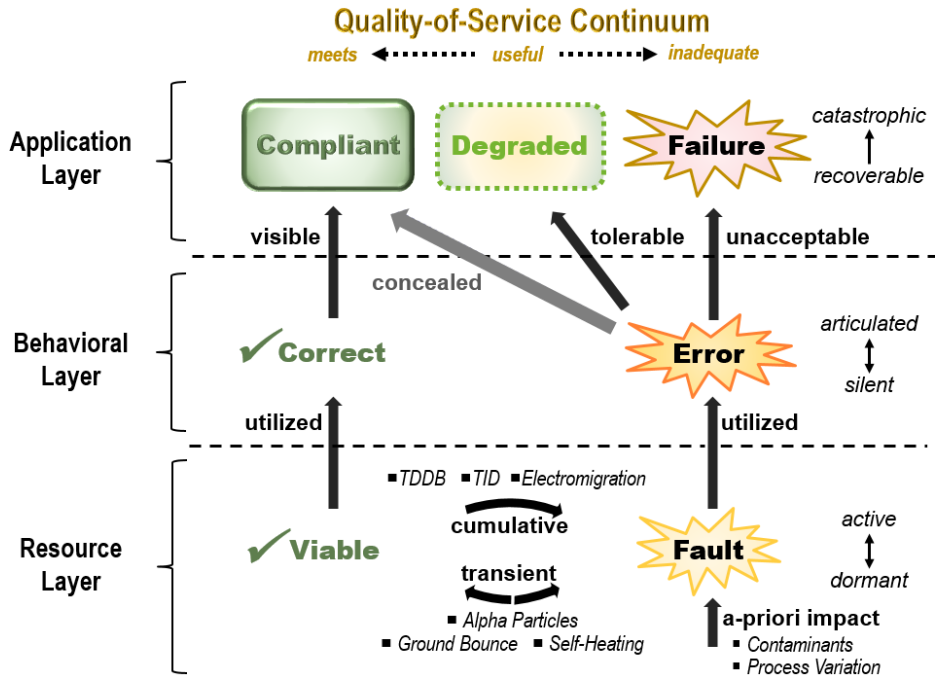

Figure 2.1: Layered model of permanent and soft error sources and impacts.

On the other hand, errors that are articulated propagate up to the Application Layer. The Application Layer depicts that correct behaviors contribute to sustenance of *Compliant* operation. Systems that are compliant throughout the mission at the application layer are deemed to be reliable. To remain completely *Compliant*, all articulated errors must be *Concealed* from the application and remain within the Behavioral Layer. Articulated errors that reach the application cause the system to have *Degraded* performance if the impact of the error can be tolerated. On the other hand, articulated errors that result in unacceptable conditions to the application incur a *Failure* condition, which might be recovered by replacing the effected cell(s) in the configurable devices, i.e., *Field-Programmable Gate Arrays (FPGAs)*, or failures may be catastrophic in *Application-Specific Integrated Circuits (ASICs)* devices [38].

2.1.2 Radiation-Induced Soft Errors

Sources of most temporary errors that impact system reliability are listed in Table 2.1. These sources are categorized based on radiation, SET characteristics, and variability, as each one can result either in soft error or hard error/failure. Ionizing radiation can have different effects on CMOS chips, and this depends on the radiation environment in which the chips are utilized. Concerning CMOS circuits, there are two broad categories of radiation effects including *Single-Event Effects (SEE)*, which are due to a single strike of a particle with high ionizing power (thereby causing instantaneous failure); and *Total Ionizing Dose (TID)* effects, which are due to the progressive accumulation of defects caused by the movement of many particles (causing long-term failure), such as electrons or protons. Radiation-induced soft errors at the sea level consist of neutrons, photons, muons (generated from the interaction of secondary cosmic effects with the atmosphere), and pions. Additionally, alpha particles, that either originate from some metals used in integrated circuit fabrication or emitted from some elements that are used in the doping of semiconductors (radioactive impurities), is also an important source of error in sensitive circuits. Among them, roughly 95% of these particles are neutrons [16, 28].

In SRAM cells, secondary particles from neutron collisions have the largest effect on Q_{crit} [28]. Overall, at sea level, both alpha-particles and neutrons are important sources to induce soft errors in CMOS SRAMs [39]. On the other hand, it was projected in [39] that cosmic ray neutron-induced soft errors demonstrates SER in latching circuits since they constitute a larger charge capacitance. However, as Q_{crit} becomes smaller with process scaling, alpha-SER increases 4-fold per generation [40]. This indicates that LET from alpha particles can generate greater than 1 MeV (equivalent to 44.5fc) by direct ionization, which is more than adequate to charge/discharge a node capacitance and flip the state of the struck node [28]. It is also reported that SEU susceptibility to low-energy protons and alpha particles from radioactive impurities is able to deposit adequate

charges by direct ionization to cause SEUs due to an intrinsic reduction in the amount of deposited charge [16, 28].

Practically, enhancing the purity of the materials reduces the SER induced from alpha particles, whereas the effects of cosmic rays can be reduced by shielding the die area [28]. Shielding provides partial masking of electrons and low energy protons. In addition, shielding a circuit by packaging to prevent SEUs can increase the SER due to packaging as packaging materials themselves can release alpha particles that induce soft errors. Likewise, long-term TID exposure can result in variations in threshold voltage, which might increase device leakage currents, and thereby increase power consumption. These are in addition to increasing the degradation of materials, leading to timing changes and thus decreased functionality. Soft error effects can be mitigated using radiation-hardened devices and shielding. However, these strategies are inadequate for robust systems with high reliability. Thus, utilizing redundancy techniques or changing the geometry of the device are still in use to improve the reliability. For example, the experimental results in [41] projected that SER for SOI is reduced 5-folds compared to substrate bulk technology. In addition, the contribution of neutron and alpha particle induced soft errors have found to be reduced roughly the same difference. Hence, the difficulty is not how to achieve a high reliability. However, achieving it with minimum area and power overheads and speed degradation is the aim. Thus, it is desired for SER mitigation schemes to allow tradeoffs between reliability and overhead of protection techniques. We identify these techniques in Section 2.3.

2.1.3 Sources of Variations

The lower portion of Table 2.1 identifies the sources of variability that impact reliability. Nanoscale devices are susceptible to process variations created by precision limitations of the manufacturing process [42].

Table 2.1: Categories of exposures facing highly-scaled CMOS logic circuit design.

Category			Source	Failure Mechanism	Effect	Error Type
Background Radiation	**Charged Particles**		Alpha particles(α)	Direct ionization	Single or Multi-Bit Upset (MBU)	Soft error
			Proton(p)	Direct ionization	Single Event Upset (SEU)	Soft error
			Beta(β) or Electron(e)	Direct ionization	SEU	Soft error
	Cosmic Rays		Photon: Gamma(γ), X-ray	Indirect ionization	SEU at sea level or Single Event latchup (SEL) in space level	Soft or hard error
			Neutron(n), heavy ions	Indirect ionization	SBU/MBUs at sea level or SEL in space level	Soft or hard error
	Secondary Radiation		Sea level muon	Direct ionization	SEU	Soft error
SET Characteristic Issues			Propagation-Induced Pulse Broadening (PIPB) or Pulse Stretching	Increase the width of a transient pulse	Increase SER	Soft error
			Pulse Quenching	Reduces the width and amplitude of a SET	Reduce SER	Soft error
Variability in Devices	**Supply Voltage Variation**	**Static Variation**	Occurs during fabrication process, such as RDF, LER, WKF, and etc.	Manufacturing imperfection	Time-dependent (increase/decrease output delay)	Bit error
		Dynamic Variation	Supply voltage variations, such as voltage drops by the effect of higher heat flux	Operating margin violation	Time-dependent	Bit error
	Oxide Thickness Variation		Occurs during fabrication process	Manufacturing imperfection	Inherent unreliability effects in a logic circuit	Timing/Bit error
	Gate Length Variation		Occurs during fabrication process	Manufacturing imperfection	Inherent unreliability effects in a logic circuit	Timing/Bit error

Sources of variability can be categorized into two groups: *intrinsic (random)* and *extrinsic (systematic)*. Process variations on the gate oxide thickness (t_{ox}) and the gate length (L_g) are the most dominant extrinsic variations, whereas variation that is intrinsic in nature results from stochastic behaviors related to device fabrication including *Random Dopant Fluctuation (RDF)* (due to random number and position of impure dopants), Linear Edge Roughness (LER), *Random Work Function fluctuation (WKF)* (due to metal gate), and *Interface Traps (ITs)* [43, 44]. Practically, to alleviate the intrinsic-parameter fluctuations, the high-κ/metal-gate technology was introduced as an effective alternative.

Considering all sources of variations, one of the most significant is RDF. Additionally, RDF is exacerbated as the device area decreases for bulk substrate devices. It was reported in [1] that the reduction in technology size accompanied by each technology generation decreases area roughly by half. Therefore, the channel dopant concentration decreases exponentially over time. Consequently, due to a few tens of dopant atoms in the channel for generations beyond 32nm, compared to thousands of dopant atoms in earlier technologies, the randomness will increase causing more variability [1]. However, in FinFET devices that have undoped channels, i.e., HK-MG Silicon-On-Insulator (SOI) developed by IBM, RDF is considered negligible and WKF, ITs, and LER become the major sources of fluctuation [45,46]. The increased occurrence of process variation results in a distribution of threshold voltage (V_{th}). As the V_{th} increases, the increase in switching time affects the delay performance of the circuit. Such variability is observed to be magnified by the continued scaling of process technology [46]. For example, the effect of RDF is magnified as the number of dopant atoms is fewer in scaled devices, as shown in Figure 2.2. This implies that the addition or deletion of just a few dopant atoms significantly alters transistor properties.

Nano-scale devices in CMOS technology require eliminating the fabrication challenges (extrinsic variation) as well as mitigating intrinsic variation effects, which are essential for the characterization of deeply-scaled CMOS planar and FinFET structures [31, 43]. In the design levels, the Process Variation Effect (PVE) and RDF predominate at the circuit-level and device-level fluctuations [47]. Although all of these factors influencing variation contribute a major role in impacting system reliability, recently, soft errors have been pronounced in the literature as predominant sources of reliability degradation, especially under the scaling impact at recent CMOS device technology nodes [16, 26, 27]. Therefore, herein we concentrate on radiation-induced transient faults and review techniques presented in the literature to immune logic circuits and systems towards soft errors. Since the characteristics of soft-errors impacting logic and memory elements differ, we review them separately in Section 2.2 and Section 2.3.

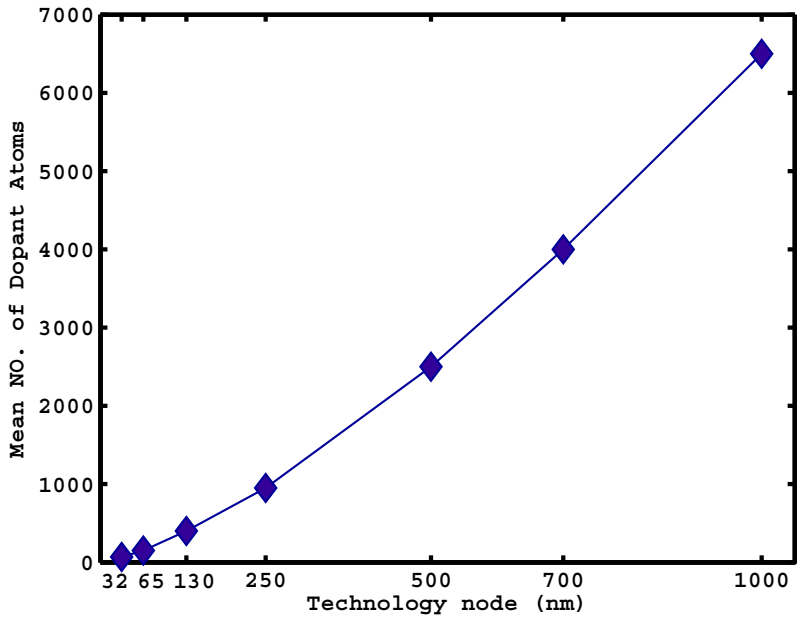

Figure 2.2: Random dopant fluctuation as a function of technology generations [1].

2.2 Soft Errors in Logic Paths

The contribution of soft errors in logic datapaths as opposed to memory elements is becoming significant as the technology node and supply voltage are scaled down, specifically at the 65nm process node and beyond [17, 27]. In the past, designers focused on soft-error induced by SEUs to protect memories rather than protecting logic paths, i.e., latches/flip-flops and combinational logic gates, where transient logic errors were considered negligible as compared to SEUs in memories [20]. However, as device shrinkage leads to a reduction in the amount of collected charge, the SER trend has been higher because of the reduced quantity of collected charge required to induce voltage and current swings in CMOS logic circuits. Therefore, as research studies and circuit designers progressively advance technology processes including shrinking the size of transistors, reducing power supply voltage, growing integration capacity, and increasing the switching

speed of charging/discharging operations, these have reduced a circuit's reliability and increased the complexity of submicron ICs [19, 20]. For instance, Dixit and Wood [10] examined SER in a microprocessor, and their experimental results revealed that SER is expected to exaggerate under technology node scaling. This causes the unprotected part of SER per microprocessor to be increased even though the SEU per bit has decreased. Therefore, leaving the SER microprocessor logic (latch-origin) unprotected as in earlier technology nodes has become reliability concern for advancing technology generations. This is due to device feature size being decreased, leading to SER in microprocessor logic becoming one of the major contributors that might cause system failure [10, 14]. Thus, as with memory elements, it has become essential to protect logic paths against soft errors [17], with emphasis on protecting the latching circuits.

2.2.1 Logic Path Masking Mechanisms

In logic paths, the propagation of a transient pulse through downstream logic might be masked by three inherent masking mechanisms. These can prevent the propagation of a spurious transient pulse along a path towards the input of a flip-flop/latch, where it may be registered to cause an SEU [48]:

1. *Logical Masking:* corresponds to cases when a transient pulse does not affect the computation in other gates along the path towards the output for a given input vector,
2. *Electrical Masking:* caused by attenuation of the glitch while passing through subsequent logic gates, and
3. *Latching-window/temporal Masking:* occurs when a generated glitch does not occur within the setup/hold time window of a flip-flop.

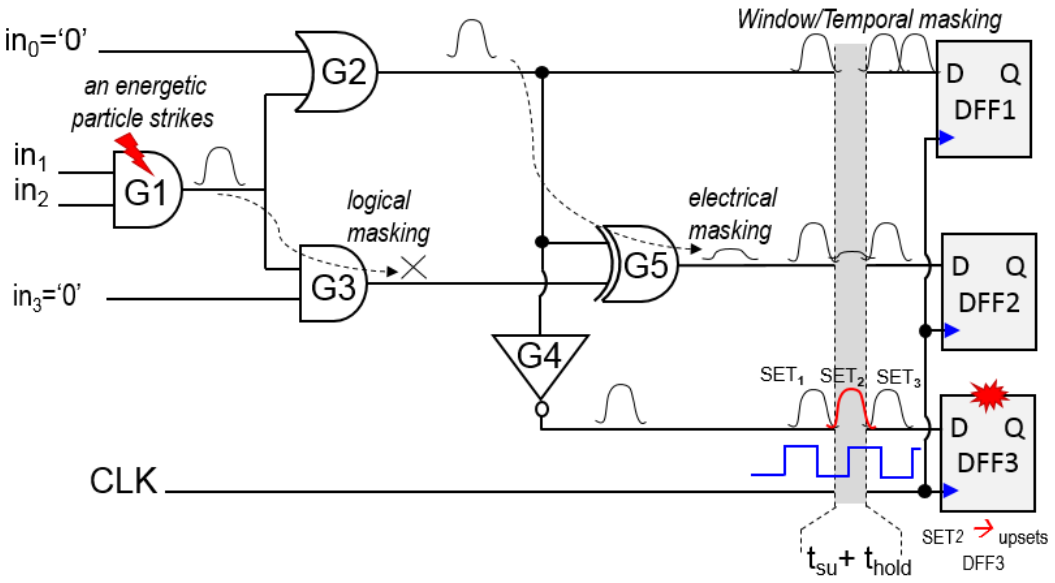

Figure 2.3: An example of logic datapath masking mechanisms.

Figure 2.3 depicts a brief example to elucidate these masking mechanisms further. As seen in the figure, the generated SET pulse at gate G3 is masked by the logical masking since the output of a AND gate is logic 0 whenever one of its inputs is 0. On the other hand, the SET pulse at gate G5 is attenuated by the electrical masking due to the assumption that its width is smaller than the transition delay time of circuit G5, however, the SET pulse maintains its amplitude while passing through gate G4 since its width exceeds the transition delay of G4, as assumed in this example. While the SET at DFF1 and DFF2 is not captured to flip the bit state and causes an upset because it is either masked by the temporal masking, i.e., does not arrive during the setup/hold time of DFF1, or it arrives at the capturing time, it was attenuated by the electrical masking, i.e., at DFF2. Thus, its amplitude is inadequate to cause an upset. Meanwhile DFF3 undergoes an upset (bit state flip) due to the transient pulse's (SET_2) arrival at the setup/hold time of DFF3.

These masking mechanisms have been considered to reduce SER in the logic datapaths significantly as compared to that of a memory cell for the same technology process [48]. In addition,

they are exploited in the SER mitigation approaches at the circuit/module-level as the means to eliminate the propagation and/or capturing of SET pulses. However, the downscaling in the transistor's structure and supply voltage reduction has reduced the benefits of the masking phenomenon and made transient faults a potential source of decreased reliability in logic paths.

It is evident that logical masking is not impacted by operation at lower voltages (technology generation) since it is circuit design dependent. For example, a 0 logic on one input of a NAND gate or a 1 logic on one input of a NOR gate prevents the illegitimate pulse from being propagated to the next logic gate, or controlled by the state of the combinational logic. On the other hand, the propagation and/or attenuation of such transient pulse depends on its duration. For example, if a SET pulse duration is roughly lower than the transition delay of a logic gate, it cannot propagate through a logic path. Whereas, if its width exceeds twice the transition delay of a logic gate, it propagates through the logic without attenuation. However, it might undergo some attenuation during its propagation if its duration is between 1-fold to 2-folds of the logic transition time [20]. Moreover, at low supply voltages, the electrical attenuation is lowered through logic datapaths since smaller transistors are faster to charge/discharge, which might result in reduced attenuation effect on the SET amplitude [10, 48]. For instance, in the earlier technologies' gate capacitances utilized large amounts of charge, and therefore, they were less prone to experience upsets, and even if an upset occurs, it was most likely to be attenuated by electrical masking due that the gate transition delay was larger than the SET pulse width [17]. Likewise, Mahatme *et al.* [27] projected that electrical masking is likely to diminish as device feature size is decreased due to a reduction in the gates' capacitance and faster switching speed of transistors. Similarly, temporal masking mechanism provides reduced benefits under technology scaling. A SET pulse might be registered in a flip-flop or latch if the logic transient time becomes shorter than the duration of the pulse width. Thus, it will no longer be masked by the temporal masking due to overlapping with the setup/hold time.

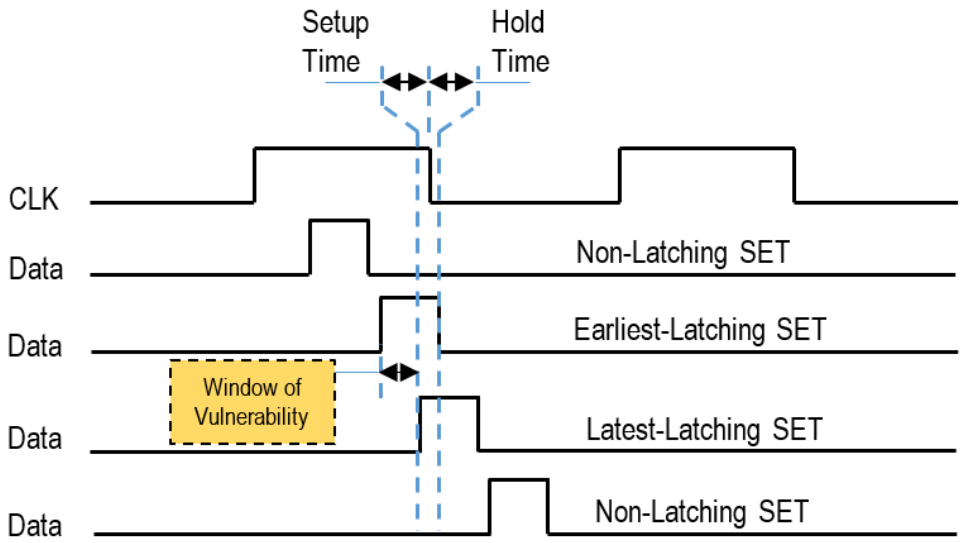

Figure 2.4: SET latching window of vulnerability [2].

Furthermore, increasing the clock frequency raises the probability in which the SET pulses overlap with the clock event, a so-called window of vulnerability [20] as shown in Figure 2.4. For instance, in [10], it is reported that the temporal masking range for latching is decreased by shrinking the transistor geometry size and/or increasing the frequency rate. Furthermore, as operating frequencies at lower supply voltage (V_{DD}) are expected to be low, it has been suggested that pipeline stages should consist of fewer gates to regain lost throughput. This will consequently lower the benefit of both logical and electrical masking.

Overall, incrementing the number of stages in a pipeline processor [48], technology node scaling, and voltage reduction can be anticipated to have detrimental impact on logic SER. Thus, it is sought to design effective soft error mitigation techniques for reliable low power applications. For instance, developing a method to select a hardening technique to protect a module of complex circuitry, so that evaluating and hardening according to the SER of combinational logic-origin or latch-origin is crucial for optimal performance [49]. In practice, designers seek the flexibility of

achieving an optimal balance between the soft-error coverage and the system requirements, with minimal overheads of power, area, and speed degradation as listed in Table 2.5.

2.2.2 SET Characterizing and Modeling

Recently, new methods have been developed to model, estimate, and mitigate SETs in digital circuits at different abstraction levels, and various SER suppression strategies have been proposed. In earlier technology generations, the largest part of soft errors causing fault in logic designs was resulting from memory-based SEUs. However, SET pulses from logic, i.e., latches and logic gates, have significantly increased when moving to advanced technology process or when increasing the integration densities [19, 20]. In particular, whether logic path originated SER or unprotected memory-originated SER will dominate the overall SER on a chip has been investigated in [10, 27]. The authors conclude that the logic SER should be emphasized for effective SER mitigation. Similarly, Mavis and Eaton [2] emphasize that transient faults, i.e., SET pulses, in logic circuits have become a serious reliability challenge under the shrinking of the MOSFET feature size.

Modeling SETs pulse width under technology node scaling has been investigated widely. In the literature, some predictions of the transient pulse width trends under technology process scaling have been conflicting [9]. In fact, the impact of technology scaling on SETs is found to be complicated by several divergent trends contributing to SET characteristics. These includes shrinking the device feature size (reduced Q_{crit}), reducing the supply voltage, higher clock rate, trends of datapath masking mechanisms (electrical and temporal masking), Propagation-Induced Pulse Broadening (PIPB) or pulse stretching, pulse quenching, n-well contact and spacing (to reduce parasitic bipolar effects), voltage/oxide thickness/gate length variation, altitude (for space applications), source of radiation/particles, and others [3, 16, 17, 26]. However, the results presented in [9] assist in elucidating some of the discrepancies in SET pulse width measurements presented by various re-

searchers in the last decade, especially the research work by Narasimham et al. [50]. In [9], the variation of transient pulse widths as a function of technology scaling was investigated. A consensus was reached through experimental verification whereby transient pulse widths in bulk 65nm CMOS technology overall was found to be reduced with technology scaling under the conditions that they evaluated, illustrated in Table 2.2.

Table 2.2: Estimated maximum SET widths based on linear energy transfer (LET) [9]; taking into concentration the effect of Propagation-Induced Pulse Broadening (PIPB) and without the effect of PIPB phenomena.

LET (MeV-cm^2/mg)	**Technology [90nm]**		**Technology [65nm]**	
	SET Width with PIPB (ps)	**SET Width w/o PIPB (ps)**	**SET Width with PIPB (ps)**	**SET Width w/o PIPB (ps)**
10	900	125	170	50
20	1200	240	175	55
30	1220	270	200	90
40	1220	360	215	125
50	1300	440	240	145
60	1400	500	250	155

The most significant outcome from Table 2.2 is that the maximum SET widths for a Linear Energy Transfer (LET) of 60 MeV-cm^2/mg are about 250 ps for 65nm, whereas they are about 1.4 ns for both 130nm and 90nm. Another crucial conclusion is that the Propagation-Induced Pulse Broadening (PIPB) also contributes a role in determining the SET pulse width. However, the broadening length that was considered in [50] was not realistic since the authors determined the longest SET pulse width by taking into account a linear chain of 1,000-inverters, whereas in real logic circuits the logic chains are relatively very short [9]. Thus, the results that are listed in the third and last column are more accurate and realistic. They demonstrate the trends of shrinking the width of SET pulse under technology generation by measuring SET pulse widths for different technologies, 130, 90, and 65nm. Clearly, the trends of transient pulse width are found to be decreased under tech-

nology scaling. A related work for determining the broadening SET pulse width in combinational logic was done in [51] using 65nm technology process from TSMC. Hamad *et al.* [51] projected that if the SET pulse width is sufficiently larger than the threshold combinational logic, it will propagate through logic paths without attenuation. For instance, their experimental results show that in a chain of six 4-input NAND gate, the width of the injected SET broadens from 100 to 178 ps when passing through combinational logic, whereas it expands to 184 ps when passing through 4-input NOR gates chain. However, this variation, i.e., PIPB, can be controlled by designing gates with balanced rising and falling times [52]. In addition, it was reported in [53] that SER of logic paths is not susceptible to the effect of PIPB phenomena except for long datapaths circuits. Meanwhile, a SET that hits the control lines have been specified as a considerable portion of the overall SER per chip.

On the other hand, parasitic bipolar amplification is found to be a major contributor that causes long widths of transient pulse in deeply-scaled CMOS technologies [3]. Amusan *et al.* [3] measured the widths of transient pulses in a 90nm technology and they report that the parasitic bipolar effect and the widths of transient pulses greatly rely on the n-well contact area and spacing. The experimental results in [3] illustrate that the widths of SET pulses can be reduced significantly by modifying the area of the n-well contact while maintaining the well size constant, shown in Figure 2.5. Similarly, Gadlage *et al.* [9] investigated SET pulse widths based on modifying the n-well contact size of the pMOS device, since parasitic bipolar amplification is more noticeable in a bulk pMOS device with an n-well and p-substrate [9]. Their experimental results show the same conclusion as the results presented in [3]. Consequently, both results in [3] and [9] confirm that layout design techniques, i.e., n-well contact area, contribute a major role in determining SET pulse width. Thus, Gadlage *et al.* [9] concluded that if the results of one work is done using a robust scheme for n-well contacts and spacing, the SET pulse widths will be decreased with technology scaling. On the other hand, if a less robust n-well contact and spacing layout scheme is employed to experimentally measure

SET pulse widths, the results will lead to the conclusion that layout n-well contact and spacing will dominate the effects of transient pulse widths.

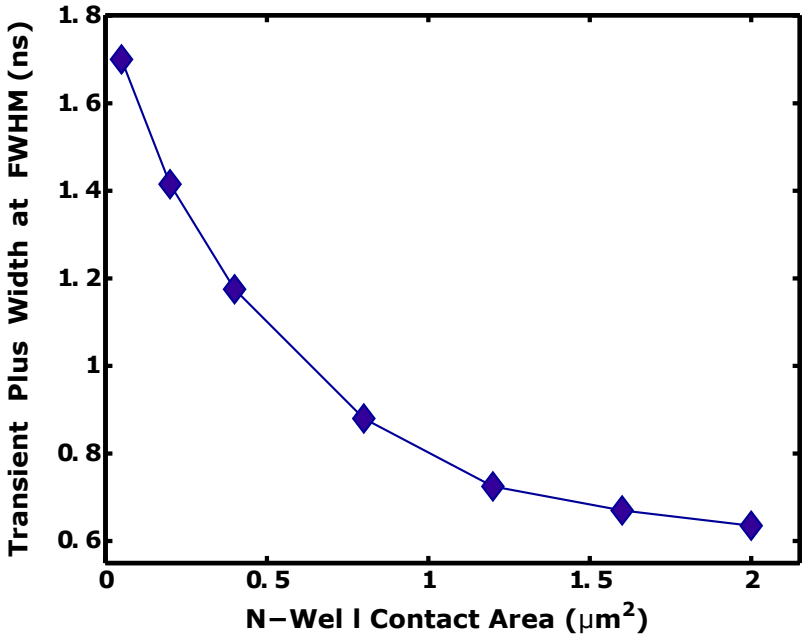

Figure 2.5: Estimated maximum SET widths based on modifying the n-well contact size of the pMOS device for 90nm technology process [3].

Overall, the variation in a SET pulse width can be either broadening its duration or attenuating its amplitude. For instance, it was predicted in [54] that, the SET pulse width is decreased as technology feature size is scaled down, where the critical transient pulse width is estimated to be 85, 70, and 55 pecs for 180, 100, and 45nm, respectively. This trend is found to be true for both bulk and SOI devices. Similarly, estimating the width of the transient pulse, induced from neutron effects, under technology node scaling was done in [55] based on changing different design factors, including: drive strength, area of the diffusion drain, fan-out, temperature, and supply voltage. Results indicated that the SER of SET in logic circuits deceases by roughly 50% as technology is scaled from 90nm to 45nm for the same circuit and clock rate. Also, the authors concluded that the main factors that should be focused on are cell type, fan-out, and supply voltage. When reducing

the drive strength or increasing fan-out, the impact of SET will decrease for the small SET pulse width (63 to 86 ps), but it will increase for the long pulse width (> 109 ps). Alternatively, supply voltage reduction will increase the rate of SET since less amount of collected charge can change the state in an output node at a logic cell.

On the other hand, since there exist several divergent trends contributing to increase or decrease SER, modeling and measuring SER is found to be complicated with advancing technology generations. Diverse analytical techniques were recently presented to assess SER in semiconductor logic circuits. For instance, Ashraf *et al.* [56] proposed a novel approach to analyze and estimate the propagation of a transient fault through downstream computing logic. The proposed algorithm utilizes machine learning techniques to precisely track the impact and propagation speed in distributed MPI applications and provide more accurate prediction, thereby preventing the erroneous outcomes/conclusions that the statistical fault injection analysis might lead to, which is based on output variation. It was stated that the introduced fault propagation framework can assist in realizing high resilience by improving vulnerability of HPC applications as it not only accurately predicts the number of corrupted memory locations into parallel MPI applications but also indicates how quickly a transient error propagates into the application's state. Furthermore, recently, researchers have switched to using probabilistic methods for modeling SET since it provides accurate and fast measuring of SET pulse width in logic circuits under the effects of process variation. For example, Yao *et al.* [57] proposed a statistical SER analysis technique. Their experimental results emphasize that PV effects have made the analysis of SET faults more complicated as the technology process shrinks beyond 40nm. Also, it has been reported that process variations of effective channel length (L_{eff}), threshold voltage (V_{th}), and oxide thickness (T_{ox}) have significantly impacted SER by more than 50% of the static analysis result of SER, thus using only a static SER analysis methodology results in inaccurate estimation of SER in logic. Therefore, utilizing a probability distribution method to characterize SER is desirable since it takes into consideration the

effects of process variations and reduces the computational time as compared to the Monte-Carlo analysis. Similarly, a system's overall performance can be improved by shifting from deterministic methods of design towards statistical and probabilistic methods due to a reduction in transistor's variations. Thus, accurate and efficient methods for estimating SER in current technologies are sought.

In summary, numerous recent research studies emphasize that soft errors will increase and negatively impact the reliability of CMOS circuits as technology is aggressively scaled down [16, 21, 29]. This is because Single-Event Multiple Transients (SEMT) events have increased as device technology scales down. This implies that the probability of double or even multiple errors occurring simultaneously has increased [58]. For example, a single particle strike might impact double adjacent vulnerable nodes of a logic circuit as a result of shrinking the space between susceptible nodes and also reducing the node capacitance and supply voltage. Consequently, following Moore's law, it is predicted that SER per logic bit increases roughly by 8 percent per technology generation [1, 7]. This is due to the number of transistors per unit area doubling. Thus, SER per unit area continually increases even though SEU per bit has decreased [1, 16, 17, 26], shown in Figure 2.6 [1]. For instance, Intel reported worsening trends that have spanned from 0.25μm to 90nm [7].

2.3 SER Mitigation Techniques

In order to achieve high-reliability in contemporary computing systems, memory protection techniques have been a prominent development. However, such techniques alone have become increasingly inadequate to completely and efficiently protect the reliability of VLSI circuits and systems. Since memories are protected with Error-Correcting Code (ECC) and parity via information redundancy, the residual SER originates from logic circuits. In the literature, wide ranges of strategies

have been proposed to protect CMOS memory and logic circuits versus soft errors.

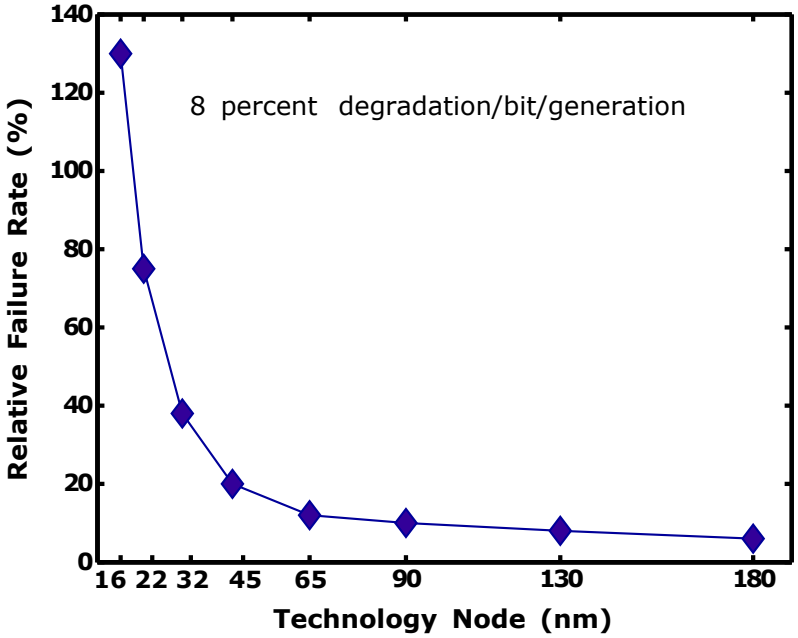

Figure 2.6: Soft error rate of a logic and memory [1].

Figure 2.7 shows a new contemporary taxonomy of soft-error mitigation spanning multiple layers of abstraction and the general resolution characteristics. We classified these techniques into three levels including Radiation-Hardening By Process and material (RHBP) techniques, i.e., Silicon-on-Insulator (SOI) at the device-level, Radiation-Hardening By Design (RHBD), i.e., spatial and temporal redundancy at the gate-level and circuit/module-level, and resilience by coding techniques, i.e., information redundancy at the system-level [16, 20, 26, 59]. Several modifications to SER mitigation schemes at these levels have been adopted. Each approach exploits a design property to achieve system constraints while accommodating some inherent challenges. Furthermore, SET hardening techniques can be categorized into three groups, including the point of SET origin (SET generation), along the datapath (SET propagation), and within the latch (SET capturing) [16, 26]. Schemes that tackle soft errors at the SET's origin concentrate on limiting the

collected charge at the sensitive node of an OFF state transistor, i.e., the drain side. Limiting the collected charge diminishes SET pulse width and reduces SET voltage amplitude [26]. Reducing SET's voltage amplitude boosts the electrical masking when a SET pulse propagates through downstream logic. Additionally, device resizing [60] for current drive is an obvious SET mitigation technique. Also, specifying the most critical logic data paths, which have nodes more susceptible to cause an error, is considered an effective concept. In particular, one of the most effective techniques for hardening SETs is inside the latch, which leverages the property that not all SET pulses will arrive at the setup/hold time of a flip-flop. Thereby, the transient pulses, which do not overlap with the window of vulnerability of a storage element, will not cause an upset [26]. Techniques that are designed based on neglecting the irrelevant SET, i.e., not captured, are profitable concepts and they have been adopted to reduce the overheads of spatial redundancy approaches. An overview to discuss various tradeoffs of each sub-category of techniques will be presented next to highlight the pros and cons of SE mitigation at each design level.

2.3.1 Device-Level Mitigation Techniques

At this level, designers concentrate on reducing the amount of collected charge by optimizing the fabrication process, i.e., using more advanced and sophisticated manufacturing processes to improve layout designs. Reducing the collected charge realizes a reduction in SET amplitude, which assists in realizing better electrical masking, whereas switching to use Silicon-on-Insulator (SOI) device structure, a semiconductor fabrication technique that IBM pioneered using silicon oxide with pure silicon to realize CMOS circuits and microchip fabrication, improves both temporal and electrical masking due to its buried oxide that reduces the amount of collected charge. Likewise, CMOS SOI-based device structure is more resilient to multiple-cell upset than the bulk structure due to its inherent resistance to the effects of SEMT [16]. Meanwhile, SOI devices are more susceptible to PIPB phenomenon, i.e., the change of pulse characteristics during propagation, as

precipitated by body potential fluctuations, which results from having a floating body.

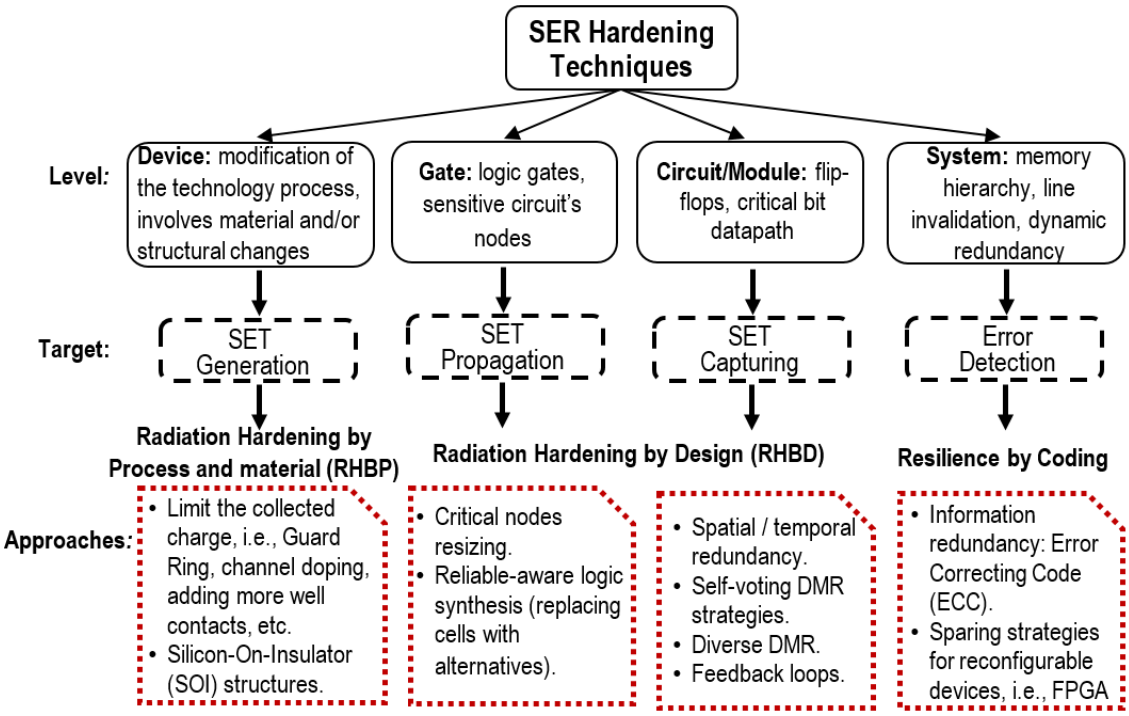

Figure 2.7: Taxonomy of design levels for soft-error mitigation schemes.

Other techniques that also exploit the inherent masking mechanisms are guard-drain and guard-ring for P-hit and N-hit mitigation, respectively [61]. Additionally, modifying the channel doping concentration of the well itself, or adding more well contacts helps to eliminate or reduce soft error effects. Moreover, layout topology participates in determining the pulse quenching phenomenon that leads to reducing the width of the SET pulse and SER at the device-level. For instance, the experimental results in [62] conclude that benefits from pulse quenching are increasingly pronounced in the common n-well design as compared to the separate n-well design of an inverter chain. This confirms that layout design techniques, i.e., n-well contact area, contribute a major role in determining SET pulse width, and thus reducing SER in CMOS circuits [62]. Furthermore, Lee *et al.* [63] introduced a new layout topology called *Layout Design through Error-Aware Transistor Positioning (LEAP)* that can be used to improve soft-error resilience in terms of single event mul-

tiple upsets for sequential elements, such as latches and flip-flops. The principle of the proposed design is based on physically combining or placing the layout of susceptible contact nodes of any two OFF state transistors with an ON state transistor so that reducing the collected charge, and thus, compacts MBU malfunctions. The design principle of LEAP was tested on DICE-based and conventional flip-flop circuits, and the results showed that the LEAP-DICE achieves 5 and 2,000 orders of magnitude soft error resilience improvement, respectively. In addition, the LEAP-DICE flip-flop incurs negligible power and delay overheads compared to a conventional DICE flip-flop due to the schematic structure and the operating behavior remain without any modification. However, since the principle of LEAP design utilizes more complicated intra-cell routing, it incurs roughly a 40% area overhead.

Similarly, adding capacitance in the feedback loop, between upset and recovery, has been successfully used to improve the speed penalty of SEU suppression, especially for the write operation of SRAM, and to improve the effectiveness of feedback resistance [64]. However, as we move towards advanced technology scaling processes, utilizing more robust fabrication processes adds more complexity and increases the production cost [14, 29].

2.3.2 Gate-Level Mitigation Techniques

Herein, soft error masking techniques rely on protecting the vulnerable nodes in logic paths. These schemes prevent SETs from being propagated through the datapath, including Cascade Voltage Switch Logic gate (CVSL) [65], critical node ranking, or reliable-aware logic synthesis, i.e., replacing cells with alternatives [66]. The drawbacks of these approaches are that they impact the system's speed performance more significantly than redundancy approaches at the circuit-level, thus they might not be attractive for high speed reliable systems designers. In addition, the use of gate resizing and gate cloning have been introduced as alternative to reduce the overheads of

spatial redundancy, i.e., TMR, and improve SET masking of logic paths in CMOS technology. These techniques search through a logic path to evaluate the most susceptible nodes of a logic circuit. Once the susceptible nodes are determined, the logic gates are resized by increasing their capacitance, so that Q_{crit} is increased and the prospect of SET occurrence is reduced. However, resizing the sensitive gate nodes does not guarantee a reduced SET pulse width [52]. Furthermore, it is costly and quite difficult to resize each sensitive individual node in designs consisting of over 1 billion transistors. Additionally, the improvement of hardening techniques at this level may be partial at-best, thus considering some redundancy at a higher level of abstraction is required for complete and cost-effective SEU handling [29].

2.3.3 Circuit-Level Mitigation Techniques

Circuit-level SER mitigation techniques have been introduced as intriguing alternatives to reduce the design complexity and fabrication cost of SER suppression schemes designed at the device-level [67]. In the literature, numerous design techniques have been introduced at the circuit-level to mitigate soft errors in logic computing circuits and systems, such as spatial, temporal, and hybrid redundancy approaches. The positive aspect of these SER handling techniques is the ability to protect both sequential and combinational logic circuits, aside from mitigating both SET and SEU errors by utilizing a single scheme. However, they involve high area overhead, power consumption, and/or speed degradation. Thus, tradeoffs between these performance penalties and the fault masking coverage are sought, so that hardening only the sequential logic portion, i.e., latches and flip-flops, which contributes the major portion of SER in logic circuits compare to that of combinational logic [21, 68, 69]. This significantly reduces area overhead, improves circuit and system reliability, and lowers production cost. In particular, protecting latches against SEU is a prominent research focus, where replacing a regular flip-flop (unprotected) by a SEU-immune flip-flop improves soft-error by roughly 10-folds [1, 70]. Next, we discuss the state of the art

of radiation-hardening latching circuits and also the conventional redundancy-based mitigation approaches that are introduced to address the effects of soft error.

2.3.3.1 State of the Art Redundancy-based Techniques

Over the years, various techniques have been proposed to increase the radiation-tolerance of CMOS-based latches and FFs, of which the most recent related approaches are highlighted in Table 2.3. Here, the mean idea of the previously presented approaches is either increasing the feedback loop delay to introduce longer delay than a potential transient glitch that might impact a vulnerable node of a circuit or utilizing the module redundancy by duplicating/triplicating the circuit and employing a majority voter circuit to determine the final masked output. Notice that the condition of a transient pulse (SET) to upset a latch or flip-flop circuit is that its duration has to be larger than the feedback loop delay of a latching circuit [71]. The comparison among various approaches is conducted based on the design-level of protection techniques, considering the overheads for hardening, fault-masking coverage, and design challenges as listed in Table 2.3.

Table 2.3: Comparison of selected circuit-level protection techniques.

Design	**SE Prediction Technique**	**Resilience**	**Feature**	**Challenges**
Mitra et al. [15] (BISER)	Module Redundancy	SEU	- double conventional latches - comparator - weak keeper circuit	- high area - vulnerable to DNU
Seifert et al. [72] (RCC)	Feedback Loops Cross-coupled Inverters	SEU	- increase Q_{crit} - dynamic process for charge collecting	- high cost - impact by voltage and technology scaling
Katsarou et al. [58] (DNUT-DICE)	Module Redundancy Feedback Loops	DNU	- 2 DICE-based latches - comparator	- high area - high power
Yan et al. [68] (DNURL)	Module Redundancy Feedback Loops	DNU	- 3 SEU-immune latches - 9 C-element	- high area - high power
Watkins et al. [73] (HRDNUT and TNU-latch)	Module Redundancy Feedback Loops	DNU and TNU	- triple conventional latches - 3 C-element	- high area-power - high delay (D-to-Q)

There are several schemes that have been developed to increase the robustness of the latching circuits, so that they become resistant to radiation-induced transient errors. For instance, *Dual Interlocked-storage Cell (DICE)* [74] is one of prevalent designs for hardening against SEUs inside latches due to its superior resilience of fault masking coverage and modest performance penalties. The advantage of DICE is that it utilizes only 12 transistors as compared to 12-16 transistors in a redundant memory cell (hardened cell against SEU) [22]. It utilizes feedback transistors and advanced layout design techniques that isolate susceptible nodes. Thus, it has demonstrated reduced area, delay, and power consumption compared to spatial redundancy-based techniques, however it is not able to tolerate multiple node upsets. Recent investigations of radiation-based soft errors in hardened latches have shown that even though the DICE cell can be an efficient concept to suppress SEU inside latches, it can be very susceptible to upsets when both nodes controlled by the feedback process get struck and change the state of the deposited charge, especially under technology node scaling as the cell is packed very tightly [22]. For instance, at 130nm technology node, the DICE cell is able to deliver about 10 times improvement [74] of circuit robustness over the conventional flip-flop. However, it exhibits reduced benefits under technology scaling, i.e., 1.4 times improvement for 40nm technology [75].

Jagannathan *et al.* [71] presented a radiation-hardening approach to protect latches/flip-flops. The architecture and the principle of the proposed latch is similar to that of a DICE latch where it has four storage nodes to store a single data bit and its complementary, however, the interconnection topology is different from that of the DICE latch. More importantly, the Quatro latch has larger feedback loop delay, and thus, it realizes a higher immunity to SETs. Moreover, since the Quatro latch has less sensitive nodes and higher critical charge, Q_{crit}, it exhibits higher robustness to single-event induced multiple node upsets than the DICE latch [71]. For instance, the experimental results of [71] demonstrate that the Quatro latch achieves higher SER tolerance than both the conventional and DICE latches, more than 10 and 3 times improvement, respectively. In addition,

it outperforms the DICE latch by consuming reduced area and power consumption, about 30% and 40%, respectively. However, it still does not realize a complete immunity to multiple node upset, as this issue is a serious reliability concern for large-scale logic designs that are packing tightly with enormous integration density.

Built-In Soft-Error Resilience (BISER) [15] design introduces a well-known hardening technique that uses two latches in parallel along with a comparator to mask transient faults. In addition, a weak keeper buffer circuit is connected to the output node in order to provide the correct output in an event of upset when the latches' outputs mismatch, as the output of C-element is floating (high impedance state) during the mismatch of the internal latches' outputs. Likewise, the other benefit from using the weak keeper circuit is to fight the leakage current when the output of C-element floats due to both the pull-up PMOS device and pull-down NMOS device turned OFF. Considering soft errors in latching circuits only, the presented BISER technique improves chip-level SER by 10-folds at a power consumption penalty of 10.3%. However, the BISER design is only able to tolerate SEUs, and therefore, can not be considered as an efficient solution to be embedded in large-scale logic paths of high-reliability systems. A similar approach is employed in [58] to develop a DNU-tolerant DICE-based (DNUT-DICE) latch using two SEU-tolerant DICE latch circuits. The proposed DICE-based latch achieves an effective soft-error masking coverage for both single and double node upsets. However, this is achieved at the cost of increased area overhead and power consumption as listed in Table 2.3.

Furthermore, numerous modifications have been carried out to increase the immunity of DICE latch against soft errors, so that it can concurrently tolerate DNU. The primary concept of the previously presented approaches is either increasing the feedback loop delay to introduce longer delay than a potential transient glitch that might hit a node of a DICE latch, or isolating the adjacent susceptible nodes using advanced layout design techniques. D'Alessio *et al.* [76] presented a Transistor DICE (TDICE) design by adding four NMOS transistors that are functioning only when

the cell is written, whereas they are in OFF state during the hold mode, thereby enabling the proposed latch to achieve high speed performance. It was reported that the TDICE design assists in improving SEU resilience by three orders of magnitude of the critical charge for vulnerable nodes. However, the drawback of the presented TDICE circuit is that it is still susceptible to DNU.

To address vulnerability of the TDICE design, Wang *et al.* [77] introduced a DICE-based latch with feedback transistors. Similar to [76], the proposed latch requires four extra devices (two PMOS and two NMOS transistors), which are turned ON during the write access operations and turned OFF during the hold mode, to effectively tackle the impediment of DICE design and reduce the probability of DNU simultaneously, thus its robustness is extremely improved. However, this results in increased delay and power consumption since transistors were added in the critical path. Another disadvantage of the presented DICE latch with feedback transistors is that it has a single sensitive pair to cause DNU. This vulnerability can be alleviated by isolating the sensitive drain area occupied by this node pair on the cell's layout. The area and power consumption penalties for the DICE latch with feedback transistors are estimated at 33% and 26% compared to conventional DICE cell, which incurs a roughly 5% speed degradation.

A novel mitigation approach, namely Reinforcing Charge Collection (RCC) was proposed in [72]. The RCC technique is based on adding some redundancy, i.e., dummy gates, to make both the OFF and ON devices in each inverter in the cross-coupled inverters collect the generated charge via charge sharing mechanism, leading to increasing the critical amount of charge at a vulnerable node. Therefore, upsetting the stored state becomes difficult since the generated charge will diffuse. The RCC hardening technique is considered an efficient mitigation scheme that only requires less than 20% of area overhead and 28% of power consumption, while impacting speed performance about of 6.4%. However, it is inadequate to be implemented in space applications, aside from its high cost and reduced fault masking resilience under technology and voltage scaling.

Recently, a DNU-Resilient Latch (DNURL) design was introduced in [73] consisting of triple SEU-Resilient Cells, each of which includes three 2-input Muller C-element circuits. The DNURL achieves a significant reduction in latching delay that makes it a promising candidate for high-speed applications. However, these advantages are achieved at the cost of high area overhead and power consumption. Thus, as observed in Table 2.3, all previously presented circuit-level SEU-immune techniques either impose significant area/energy overheads or cannot tolerate DNU. Therefore, cost-efficient and double/multiple node upset tolerant approaches that incur reduced area-power overhead and circuit performance degradation are sought.

2.3.3.2 The Conventional Redundancy-based Techniques

The soft error resilience of conventional techniques can be obtained by applying four redundant strategies: spatial redundancy approaches such as Triple Module Redundancy (TMR), temporal redundancy approaches such as clock shadow latches, self-voting redundant system strategies (combining the spatial and temporal redundancy approaches), and diverse dual module redundancy, such as spatial redundancy with diversity. Redundancy techniques are effective to mitigate soft errors inside the latch and achieve a high level of protection by suppressing SEUs. They are utilized for designing mission-critical dependable systems so that achieving high reliability and availability while directly countering beneficial characteristics such as: low power, high speed, and minimal layout area. Thus, the challenge is to implement them with less penalties since they incur significant power and area overheads. Note that information redundancy schemes, such as Cyclic Redundancy Check (CRC) codes, are also used for fault handling but at the system-level and they will be discussed briefly in Section 2.3.4.

Most of the redundancy techniques operate based on voting mechanisms. Figure 2.8(a) depicts a conventional 3-input majority voter circuit. The output changes state when two out of the three

inputs change [4]. A transient glitch on any one input will be rejected by the voter, considering the other two inputs are correct. However, a transient pulse occurring within the voter can incur a momentary output glitch that might be corrected in the next clock cycle. Thus, no permanent SEU occurs. Figure 2.8(b) shows a circuit implementation of the self-voter, which is a 3-input majority voter configured to compare two inputs and the feedback of the output. The output of a self-voter changes state to a logical 1 when both its inputs are high, and becomes a logical 0 when the external inputs are low. In case of the two external inputs mismatch, output remains unchanged and thus prevents a SET from being propagated to the next stage [4]. Thus, the self-voter circuit is a state-holding element that behaves as a redundant copy of a latch or a flip-flop. Similar to the majority voter, a transient pulse occurring inside of the self-voter circuit results in a momentary glitch on the voter output. This is due to its input driven by a redundant copy of the same value. The self-voting circuit is also commonly referred to in the literature as a Muller C-element circuit.

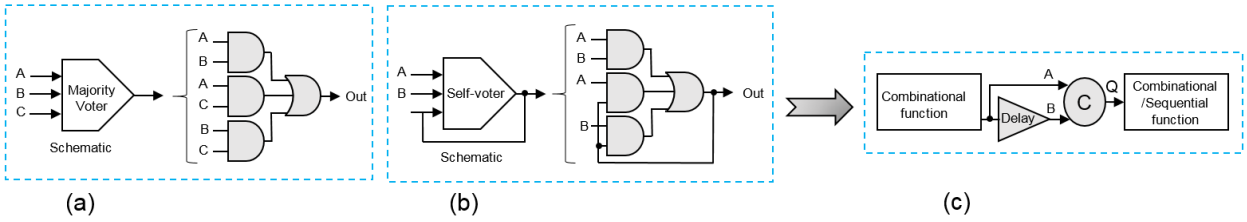

Figure 2.8: Voting Circuits; (a) Majority voter circuit, (b) Self-voter circuit [4], (c) C-element in a temporal filter structure [5].

The major benefit of a C-element design is its simplicity. A C-element maintains its output unchanged if the inputs differ, holds the previous state, or behaves as an inverter when the inputs are the same [5]. The C-element might be utilized as an evaluator between two components, as shown in Figure 2.8(c). A transient pulse that occurs at the first combinational logic, results in a discrepancy between nodes A and B, and thus, it will be masked by the C-element as long as its duration is less than the delay of the buffer circuit. On the other hand, if the width of the SET pulse

is adequately large, it will pass through the C-element to the next combinational logic. This might cause an upset if the SET pulse is captured by a storage element. The delay of the buffer circuit depends on the width of the SET being rejected. As reported in [9], there has been some conflict in the expected SET pulse widths for recent technology nodes as discussed in Section 2.2. The C-element is used to mitigate any SET with a pulse width shorter than the considered delay, by using a small area overhead. In contrast, its drawbacks are that it impacts system performance and requires some modifications for the standard library cell or design flow, besides the circuit fails if the filter or flip-flop is hit. Next, we discuss the four categories of redundancy-based techniques that are employed to protect the mission-critical applications.

Spatial Redundancy Strategies: In the literature, numerous spatial redundancy techniques have been proposed to protect CMOS circuits against soft error effects. Modular redundancy is one of the most commonly used techniques where the design is protected by replicating the circuit N times, *N-Module Redundancy (NMR)*, and extra logic is included for error detection and correction. *Dual Modular Redundancy (DMR)* is configured by duplicating the module, i.e., N equals two. In DMR, an error is detected if there exists a discrepancy between the outputs of identical modules. Conventional DMR is used to detect when a transient or upset error occurs, without error correction, since the comparator circuit is unable to identify which module is error-free [78]. *Triple Module redundancy (TMR)* is constructed by using three identical modules functioning in parallel, i.e., N equals three. Both DMR and comparison or TMR and majority voter are effective techniques to mitigate SER induced by a single SET or SEU in a logic circuit. The interest of spatial redundancy techniques are their resilience and the ability of employing them in most designs [20]. TMR is a typical example of a spatial redundancy approach to mask soft errors in logic paths, whereby the circuit or the logic datapath is triplicated and a majority voter decides the final output of the circuit [26]. Figure 2.9 depicts a simple example of TMR. Triple modular redundancy circuits are effective only in mitigating a single SET that strikes one of the modules,

considering the other two modules are error free. This indicates that TMR circuits are unable to tolerate multiple upsets that influence multiple redundant modules simultaneously.

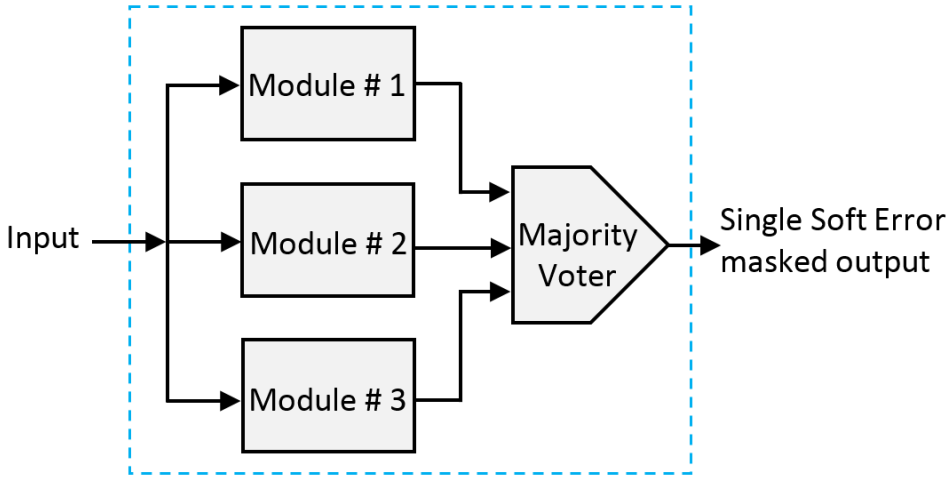

Figure 2.9: Triple Module Redundancy (TMR) approach.

Spatial redundancy is typically utilized in applications that operate in harsh environments to ensure system operation even in unforseen circumstances, such as autonomous vehicles and satellites [79–81], nuclear reactors [82], and deep space systems [83]. It has also been employed in commercial systems such as High-Performance Computing applications [84] where significant increase in computing node availability is sought. Utilization of compute-node level redundancy at the processor, memory module, and network interface can improve reliability by a factor of 100-folds to 100,000-folds [84]. This is because TMR realizes 100% fault resilience coverage for faults in single module simultaneously, compared to the simplex arrangement (unprotected design). However, this is at the cost of incurring roughly two orders of magnitude for area and energy overheads [20], which limits the usage of spatial techniques in applications with tight energy budget. Thus, TMR is suitable for high performance and mission-critical applications which can accommodate its inherent area and energy overheads.

Temporal Redundancy Strategies: In the earlier temporal redundancy approaches, input vectors are repeated for the same circuit instance to compare for discrepancy. If there is a mismatch between both execution results, then a fault occurs and the system needs to rollback and repeat operation. This scheme is immune only to transient faults, and in the best scenario, it degrades the system's goodput, or correct throughput, by 50% as each circuit is checking half the time. An evolved temporal sampling methodology, associated with minimum overheads of area, power, and performance degradation, was proposed by Mavis and Eaton [2]. They describe how the new circuit design methodology can totally eliminate all soft errors, i.e., foreseeable SETs and SEUs, in any synchronous microcircuit. Adaptions and modifications on the basic temporal sampling approach had been done to construct practical circuit embodiments that can be utilized for space and terrestrial sub-micrometer electronics. The key contribution is the ability of implementing the proposed approach in combinational logic preceding a latch or utilizing it to protect the latch itself. In this approach, data for the same combinational logic can be sampled at three distinct instances to construct a voting arrangement while using a simplex instance of the datapath. As shown in the design of Figure 2.10 [2], the data is captured at three different time instances (T1, T2, and T3) by using three identical registers (flip-flops) triggered by three different clock signals (CLK1, CLK2, and CLK3). The relative latency between the clock signals is employed such that they are delayed by a phase shift (Φ_1 and Φ_2), which can be selected depending on the SET pulse width coverage and the propagation delay time of the circuit, so that the same data from the previous stage is latched in the registers. A majority voter is employed to determine the final output, so that when a soft error occurs in the datapath logic it will be stored in one of the flip-flops, and it should be rejected by the voter. The signal $Data_{in}$ represents the coming data from the previous datapath logic, whereas the final output signal represents the data sent to the next stage. To ensure that the same data is stored in all the registers, temporal redundancy should be constrained by timing constraints, where there is a phase shift Φ_1 between CLK1 and CLK2 and Φ_2 between CLK1 and CLK3. The timing constraints are presented in [85], so that the legitimate data is captured in the

registers. Similar to the TMR arrangement, the temporal redundancy approach achieves complete SEU masking in a single module, whereas it fails in the rare case when two out of three registers are impacted concurrently. This is due to the result of majority voting.

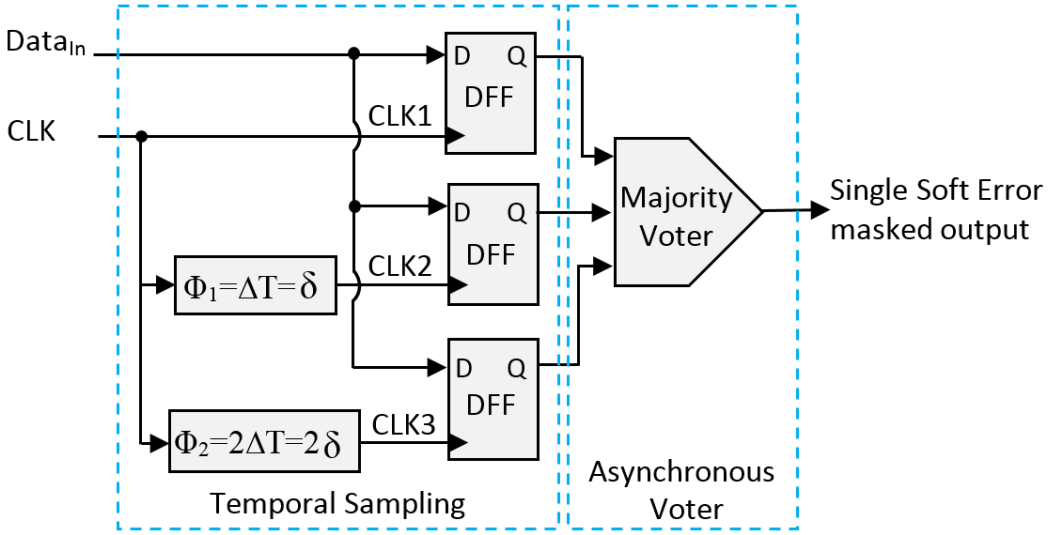

Figure 2.10: Temporal redundancy approach [2].

Avirneni and Somani [85] employed the temporal redundancy approach for mitigating soft error in a two stage pipeline processor. The authors proposed two approaches, called Soft Error Mitigation (SEM) and Soft and Timing Error Mitigation (STEM). The latter could be used to detect and correct the timing errors that occur when operating at very high frequencies, overclocking, in addition to detecting and correcting soft errors. Both approaches provide 100% fault coverage and since the proposed schemes are designed such that they do not cost any reduction in the system's overall performance during the correct operations, i.e., error-free operations, they offer better performance. In terms of performance improvement, SEM approach attains 26.58% on average compared to a conventional TMR approach, while STEM outperforms SEM by 27.42%. In [86], a heterogeneous spatial and temporal redundant FPGA-based system is designed and implemented by exploiting the linear transform of pipelineable applications. The proposed schemes are based on *Concurrent*

Error Detection (CED). In the *heterogeneous CED (hCED)* based spatial redundancy, the redundant module is used as a checker for the original module and since it only operates as an error detection circuit, its size can be reduced, which is a major concept of the hCED scheme. Alternatively, hCED based temporal redundancy utilizes the original module as the genuine module and also as the checker module by switching between them alternately. The major drawback of hCED is that it is considered an application dependent scheme since it needs a condition to identify a block as fault-free. Additionally, it can be used for error detection only (unable for correction).

Self-Voting Dual Module Redundancy Strategies: As shown in Figure 2.11, combining the DMR with the self-voting provides less area overhead compared to conventional TMR and less speed degradation (by the width of SET being mitigated) compared to temporal redundancy. The performance degradation is half that of temporal redundancy [4]. Self-voting DMR cannot tolerate multiple SETs similar to TMR. While achieving soft error masking equal to TMR, self-voting DMR must mask the effect of SET on data inputs, clock signals, and storage elements, i.e., flip-flops. In resilient systems for space or nuclear applications, the voter circuit should be protected against soft errors. Figure 2.11 shows an example of protecting an edge-triggered flip-flop besides the protection of the voter circuit, where two majority voters are used at the register output to filter out any SEU. At the temporal sampling phase (see Figure 2.11), register R1 and R2 (DFF1 and DFF2) receive their inputs directly from the datapath logic, while register R3 receives its input from a self-voter that votes on two redundant combinational logic circuits to produce its output. Legitimate data is captured by the third flip-flop as long as the two external inputs match, in the absence of SET. On the other hand, if there exists a present mismatch between the external inputs, this might result in upsetting the third flip-flop. For instance, in case the previous state of the self-voter circuit is a 1 logic and a new data, which needs to be stored in the flip-flops, is a 0 logic, and if one of the redundant modules is hit, and a SET generates and reaches the self-voter circuit, this results in a mismatch, and thus the output of the self-voter circuit will be unchanged during the

transient pulse duration. If the SET pulse overlaps with the setup/hold time of the third flip-flop, it will cause an upset/disrupt the third flip-flop's state since a 1 logic is captured by the third flip-flop instead of a 0 logic. To address this problem, the third flip-flop should be triggered by a clock cycle delayed by the width of the SET pulse. On the other hand, if energetic particles hit the self-voter circuit's output during the latching cycle of the third flop-flop, this will result in an SEU in the third flop-flop. However, the data capturing phase will not fail as long as the other flip-flops registered the correct data. The main advantage for utilizing the self-voting DMR approach is that it requires the duplication of the datapath logic instead of being triplicated as in conventional TMR, while the sequential part remains triplicated.

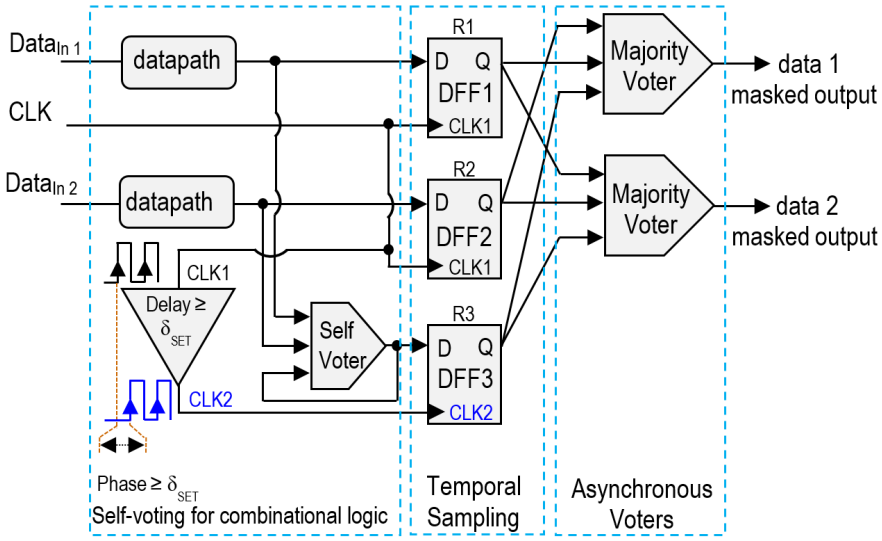

Figure 2.11: Self-voting DMR logic circuit [4].

Based on the experimental results in [4], self-voting DMR provides a performance compromise between the temporal and spatial (TMR) redundancy approaches, where it consumes less area (by one redundant combinational logic) than TMR and provides half the speed degradation of the temporal redundancy scheme. However, for designs that operate with low speed, i.e., the critical delay is large ($> 20ns$) and the speed performance of self-voting DMR becomes comparable to

the TMR [4]. Overall, the delay of TMR, temporal redundancy, and self-voting DMR are given by:

$$\delta_{critical} = \max_{1 \leq i \leq n} (\delta_i) \tag{2.1}$$

$$\delta_{TMR} = \delta_{critical} + \delta_{voter} \tag{2.2}$$

$$\delta_{Temp} = \delta_{datapath} + \delta_{voter} + 2 * \delta_{SET} \tag{2.3}$$

$$\delta_{SVDMR} = \delta_{critical} + 2 * \delta_{voter} + \delta_{SET} \tag{2.4}$$

Where $\delta_{critical}$, δ_{voter}, and δ_{SET} represent the critical path delay of the datapath logic, delay of the voting logic, and delay of transient pulse width, respectively. $2 * \delta_{voter}$ is required due to one majority voter and one self-voter are located in the longest critical datapath, whereas $2 * \delta_{SET}$ is required to ensure that the legitimate data is captured at the registers. On the other hand, if A_{logic} represents the required logic area of a design and A_{comb}, A_{seq}, A_{voter}, and A_{SVDMR} represents combinational logic, sequential logic, voter area, and self-voting DMR design area, respectively then the area overhead for the redundant systems are given by:

$$A_{logic} = A_{cobm} + A_{seq} \tag{2.5}$$

$$A_{TMR} = 2 * A_{logic} + A_{voter} \tag{2.6}$$

$$A_{Temp} = 2 * A_{seq} + A_{voter} \tag{2.7}$$

$$A_{SVDMR} = A_{comb} + 2 * A_{seq} + 2 * A_{voter} \tag{2.8}$$

Furthermore, area overhead depends on the combinational and register logic ratio in a design. Overall, self-voting DMR provides $10 - 24\%$ area improvement over the TMR implementation, based on different combinational logic ratios in the design, low, medium, and high [4]. Generally, the higher the combinational ratio, the higher area improvement achieved. Practically, it is possible to combine self-voting DMR with TMR for optimizing large chip designs against SETs, where TMR is advantageous for critical datapath pipeline stages, whereas DMR is convenient for reducing area and energy overheads on pipeline stages with non-critical datapaths [4].

Diverse Dual Module Redundancy: The conventional DMR scheme is used to detect the occurrence of a soft error, without correction, since the comparator circuit is unable to identify which module is error-free [87]. Recently, the DMR concept has been extended to provide low cost soft error rejection by monitoring the mismatch of the output, called *Diverse DMR (DDMR)*. Design diversity helps to prevent so-called *Common Mode Failures (CMFs)* [88]. By leveraging design diversity among the modules in spacial redundancy, different outputs are produced under CMF such that an error is detectable [89,90]. Also, design diversity is employed in temporal redundancy to detect permanent faults by performing the same operation redundantly by alternating physical circuits [91]. The main concept of DDMR is to use diversity with pattern mismatch, where two different structures with equivalent functionality utilized in parallel with an error location and masking circuit.

DDMR requires a discrimination circuit to be designed for each module type. When an error occurs, either structure produces a mismatch for one sample or pattern of samples and the error location and masking circuit monitors the output to determine which module structure is error-free [6]. Thus, the error pattern produced by module1 and module2 must be mutually exclusive. By rec-

ognizing the module that has a pattern of error when soft error hits, the output of the error-free module is used. The design drawback of this approach is to identify a pair of module implementations that can produce error patterns when a soft error occurs. Besides, a simple logic circuit must be used for error location and masking that only inspects the module outputs. Likewise, the redundant module should not be exceedingly larger than the unprotected module. However, *Digital Signal Processing (DSP)* is a promising field for DDMR due to the fact that DSP modules have alternative implementations in addition to consuming significant area and energy [87]. Arithmetic circuits are another identified field for utilizing DDMR.

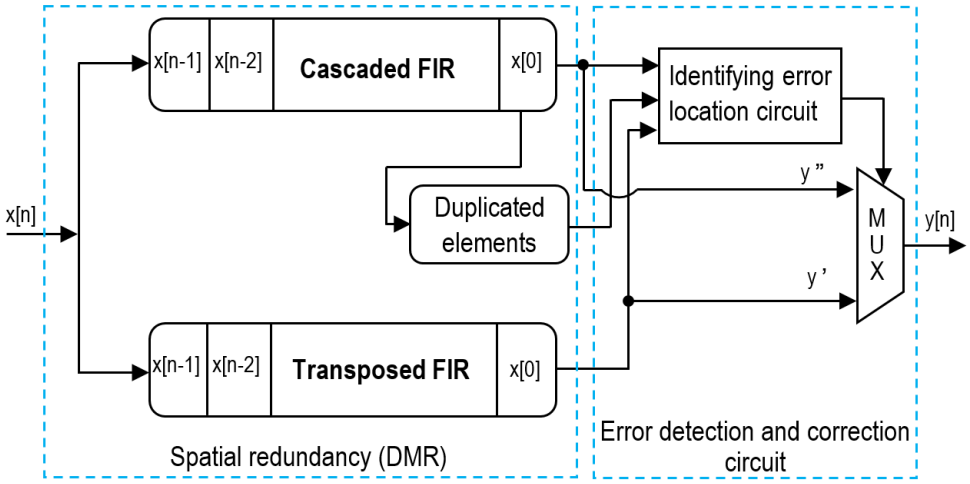

Figure 2.12: Structural DMR design for FIR filter [6].

Figure 2.12 depicts a novel technique, namely structural DMR, of utilizing DDMR to protect the *Finite Impulse Response (FIR)* filter by using two different module structures, transpose of direct and mixed transpose-cascade structures [6]. Herein, the design diversity is used to identify and mask a soft error and thus FIR filters are constructed such that they provide different error patterns at the output. A single error in the transpose filter's output, i.e., $\acute{y}$, occurs in an upset event, whereas an SEU in any stage (except the final one) of the cascade filter produces consecutive multiple errors

in output, i.e., $\acute{y}$. A duplicated element is used to duplicate the cascade filter's final stage, while an error correction (detecting and filtering out) circuit is used to check for consecutive mismatches between the filters outputs. When double or multiple consecutive mismatches are observed, the cascade filter must become upset, and thus, output of the transpose filter is selected as the final output. When a single mismatch is observed, either the error occurs in the transpose filter or in the cascade filter's final stage. Therefore, the final output is determined by comparing the duplicated element and the cascade filter's output. When a discrepancy is detected, this indicates that the error must have occurred in the cascade filter. Otherwise, the transpose filter suffers a hit and an upset occurs. On the other hand, if the correction circuit becomes upset, this only changes which filter is selected as the final output. Since the design can only tolerate a single error at one time, and it cannot occur simultaneously in both cascade and transpose filters and in the error location and correction circuit. Thus, the output of both filters must be correct. Therefore, any one of them can be used as the final output. Experimental results show that the structural DMR FIR filter consumes 2.25x area compared to the non-protected filter while TMR FIR filter requires 3.003x [6].

2.3.4 System-Level Mitigation Techniques

At the higher levels of the system stack, information-based approaches such as Error-Correcting Codes (ECCs) for memory protection can be attractive schemes to impart immunity to single-bit errors. These typically incur low area, speed, and power overheads. Multi-Cell Upset (MCU) can be suppressed by utilizing advanced coding approaches that are immune to multi-bit errors such as Cyclic Redundancy Check (CRC) codes, Reed-Solomon (RS) codes, Hamming code, etc. [16, 28, 92]. These approaches require a small number of redundant parity bits to recover in the event of an upset. However, at the chip level, memory protection techniques, or Error Correcting Code (ECC), have become insufficient for designs fabricated with advancing technology generations due to the logic soft-error, i.e., errors in latches/flip-flops and combinational logic, has

become significant contributors to increase SER at the system-level [92, 93]. This is because the spacing between the sensitive nodes is reduced with technology size scaling, which gives rise to increase the vulnerability of latches/flip-flops. Similarly, soft errors in flip-flops have become a concern due to particle strikes induced Double Node Upset (DNU) [94, 95]. Therefore, the scope of this survey is to concentrate on techniques for Radiation Hardening By Design (RHBD), thus we focus on reviewing approaches that protect the storage elements, i.e., latches/flip-flops, for reliable computing systems, while the memory cells will be reviewed briefly late on.

At the architecture-level, the effect of soft errors can be identified into two categories including: Silent Data Corruption (SDC) and Detected Unrecoverable Error (DUE). SDC errors occur when a transient fault propagates and corrupts the data, and then becomes a visible error. DUE, on the other hand, is detected by the system; however, it causes the system to freeze or crash without corrupting the user's data thereby rebooting the system is required [96]. Unlike DUE that makes the system unavailable, SDC errors result in unnoticed corrupted data. Thus, they are considered more significant in terms of reliability challenges as they might harm the system's functionality and/or user's data. This is because SDC can corrupt several components, such as registers, cache, Arithmetic Logic Unit (ALU), multiplexer (MUX), memory, etc. [97]. Error detection for arithmetic units, i.e., ALU, is achieved via information-based approaches such as Error-Correcting Codes (ECCs) [98]. Thus, the other important susceptible components are the pipeline registers. One solution to ensure data integrity of the registered states is by separating the bits of the same word. However, this scheme becomes a complicated process and might not be feasible for deeply-scaled register sets, as it impacts area and power consumption [77]. Moving forward to the architecture-level, the rate of particle induced soft errors in a processor core has become proportionally increased as the number of bits increases, which leads to impacting robustness of deeply-scaled computing systems [99]. In other words, soft errors can limit nanoscale multi-core processors used in data centers, High-Performance Computing (HPC), and mission-critical applications, such as online

banking systems [69,99]. These systems can have severe repercussions if they fail [100,101]. This is because a glitch at the circuit-level can propagate to the system-level and eventually may become a user-visible error. Therefore, efficient and practical soft-error masking techniques become highly sought after to prevent these transient faults from silently corrupting computation and logic states in registers. Achieving soft-error resilience for these resources at the circuit-level is the starting point to design reliable computing architectures. Consequently, SER of any component has to be adequately low to satisfy the system-level reliability requirement.

Similar to the circuit-level designs, the architectural state storage components are mostly hardened against soft errors using redundancy-based methods, which duplicate or triplicate the vulnerable components/circuits in the targeted system [102]. They can be protected by employing DMR to fully duplicate the pipelines, a partial redundancy using Reduced Precision Redundancy (RPR) [103], or checkpoint recovery. Both fully and partial duplication can detect and recover soft errors, whereas the latter can only be incorporated to perform recovery operations. Thus, protection of the registers at the system-level would be selectively applied to only a subset of the flip-flops, thereby maximizing the soft fault masking coverage while minimizing the performance penalties for the protection logic. Additionally, in [15], it was reported that roughly 50% of the SDC errors result from the sequential circuits, i.e., latches and flip-flops, in Intel processors. Similarly, in [104], the authors projected that the majority of the system failures were found due to the vulnerability of some important registers in the targeted LEON3 processor core. It turned out that deploying TMR approach partially to protect the CPU's most susceptible registers can achieve 99.25% of SEU masking with marginal area overhead. Furthermore, in [105], it was concluded that an important reliability improvement (masking more than 70% of soft errors) can be achieved when hardening only 10% of the sequential logic, i.e., the most susceptible registers, in an ARM Cortex-R5 CPU core. Thus, even though numerous papers in the literature discuss recent innovations for SE-mitigation in memory devices, the focus herein is on protecting the registered states

in datapaths due to its increasing importance.

On the other hand, recently, the research work in [106] showed that, even though compiler optimization has played a major contributor to realize high performance and efficiency for multi-core computing applications, their vulnerability towards transient errors has also increased, and generally the impact of vulnerability was found to be comparable to the level of code optimization. This indicates that highly-optimized code for HPC applications is more susceptible to soft errors than the unoptimized code [106]. Thus, soft error effects should be considered and required further investigation for the compiler optimization to improve soft errors masking at the system-level. Finally, at this level, an efficient soft-error suppression is achieved by concentrating on protecting storage elements, such as latches/flip-flops and memory cells. Below, how memory cells are made immune to soft errors are briefly reviewed.

2.3.4.1 Soft Error Masking for Other Components

SER has been aggravated under technology process scaling not only in logic circuits but also in memory cells. In the literature, numerous studies [10, 48, 107] have reported that SER has increased in SRAM for a fixed-sized array area whereas the device feature size has shrunk, thus the proportion of sensitive area has increased.

In modern SRAMs, the density of devices has followed Moore's law whereby recent experimental results showed that SEU susceptibility occurring from low-energy protons and alpha particles from radioactive impurities have become able to deposit sufficient charge by direct ionization to cause transient and upset glitches [16, 28, 108]. Physically, mitigating SER in memory cells involves interleaving techniques such as placing the memory layout of a one bit of many different words adjacent on a chip. Likewise, Single Error Correct-Double Error Detect (SEC-DED) codes can be effective even with Multi-Cell Upsets (MCUs) by interleaving the SRAM cells, therefore, MCUs

lead to a number of separate Single-Bit Upsets (SBUs) in multiple correction words, rather than Multi-Bit Upsets (MBUs) in a single word [109].

Unlike SRAM, DRAM-based soft errors tend to decrease with subsequent technology generations. In fact, this is due to the Q_{crit} amount being roughly constant in a DRAM cell, and thus it becomes more difficult to deposit the same amount of charge into a shrunk sensitive volume [109]. However, there exists a general agreement that MBUs should be addressed in SRAM and latching circuits since it is more susceptible as density is increased [10, 16]. At the system-level, the design process of memory cells protection is easier since it is achieved by coding. However, under the highest level, the system synchronization becomes a major problem, thereby ceasing the running task or system clock is required for synchronization [110].

2.3.4.2 Multi-Bit Upsets

Recently, as CMOS devices continue to shrink, advanced memory devices become impacted from multi-cell soft errors and thus MBU becomes a key reliability concern. An MBU occurs if two or more bits in the same word are upset by a single event at one time. For instance, spatial MBUs occur when a single particle upsets multiple bit that resides within the same physical neighborhood. Whereas temporal MBUs occur when two or more particle strikes independently upset distinct NMR instances. In contrast to SBU, MBU in memory cells cannot be resolved by single-bit ECC, thus more powerful techniques are required to suppress SER. Notice that the term Multi-Cell Upset (MCU) refers to errors that occur in different rows of an array corresponding to multiple words, thus MBUs present a serious concern as it occurs in a single word.

As the technology generations move forward, MCU exhibits a negative trend as listed in Table 2.4 [10]. Clearly, the experimental results in [10] show that SRAM-MCU has increased from 2% to 38% as the technology node is scaled from 180nm to 40nm. Single particles induced MCUs

is considered one of the crucial mechanisms that increases SER. Another source, the parasitic bipolar effects resulting from the well structure strongly affects MCUs. However, Fuketa *et al.* [111] investigated alpha-particle-induced soft error in 10T subthreshold SRAM for 65nm and they conclude that the main cause of increasing MCU in the subthreshold region is the mechanism of Single-Event Double Node Upset (SEDNU) or MBU within a device. Additionally, the variation of supply voltage and temperature also impact the rate of MBU [112, 113]. Numerous fabrication concepts are proposed to reduce/eliminate MBU, such as adding more well contacts, increasing the distance between cell to the next well contact, guard ring, etc. Similarly, SOI technology exhibits less susceptibility, and thus, it is more robust in MCU tolerance than CMOS bulk structures [16]. Furthermore, the combination of interleaving and ECC schemes are more efficient in reducing MBU rates. For example, it was stated in [114] that, MBU is reduced when the separations of data bits within each word are designed to be adequately large, whereas reducing spacing between bits in SRAMs increases the likelihood of MBU [115]. Clearly, with technology generation scaling, separating the bits of the same word becomes a complicated process and might not be feasible for deeply-scaled register sets, as it impacts area and power consumption.

Table 2.4: SBU and MCU ratio trends under technology scaling [10].

Design Rule (nm)	**SBU** %	**MCU** %
180	98	2
130	96	4
90	88	12
65	77	23
40	62	38

Considering the MBUs in current technology generations, recent studies showed that circuit-level mitigation techniques introduce reduced benefits under technology scaling [21, 58, 67]. Thus, the vulnerability of SEU-immune flip-flops increases with technology generations [58, 67, 115]. This implies that, a single particle strike that hits a sensitive node can affect more than a single storage

element at a given instance of time [58], which results in MBU in the same device. Consequently, the probability that double, or even multiple, errors occur simultaneously has become high. Therefore, soft errors in flip-flops have become a concern due to the effect of Single-Event Double Node Upset (SEDNU) or MBU within a device [21, 95]. Thus, alternative energy-efficient techniques that effectively address both SEU and MBU within a device and cross devices are sought.

Table 2.5 summarizes and provides a qualitative analysis for most of the discussed soft-error techniques. It can be seen that each approach has its inherent drawback(s). There is no optimal solution that can provide efficient performance in terms of power consumption, area overhead, speed degradation, fault coverage, and design complexity. For instance, the C-element approach is a simple design that achieves an effective reduction for area overhead, but it incurs a speed degradation, limited by the width of the SET pulse, and also needs some modification for the standard library since it is implemented at the device-level. On the other hand, the self-voting DMR approach provides a comparable area overhead to DMR, making it advantageous for the applications that can tolerate some speed degradation while achieving energy saving and handling soft errors. Meanwhile, DDMR occupies a small area compared to the conventional TMR arrangement and incurs less speed degradation than temporal redundancy approaches. These are powerful advantages, especially for DSP applications and arithmetic circuits, but it requires one to identify a pair of module implementations that can produce a pattern mismatch. Developing a new concept that can be used for a better optimization of the self-voting DMR approach or finding other promising fields that can be implemented using DDMR to protect a design against soft error effects are still challenges for voting approaches based on DMR. On the other hand, systems within iso-energy constraints can be protected by hardening only the sequential logic portion, i.e., vulnerable registers, thereby tolerating some unmasked soft errors while significantly reducing the protection cost. Thus, it is inconclusive to determine which approach can be utilized to achieve the design requirements without considering the application and properties that the design is used for.

Table 2.5: Summary of SER mitigation techniques proposed at different abstraction design-level.

Approach	Level	Scope	Beneficial Attributes	Performance Penalties and Costs
SOI structures [41]	Device	Optimizing the fabrication process, material and/or structure change.	SER is 5-fold lower than for bulk technology.	More susceptible to thermal effects and pulse broadening phenomenon.
Guard ring and guard drain [61]	Device	Limit the quantity of collected charge at the sensitive node of the transistor.	Diminishes SET pulse width (improves the temporal masking) and reduces SET voltage amplitude (boosts the electrical masking).	Increase fabrication cost under technology node scaling.
LEAP-DICE [63]	Device	Physically placing the layout of an invulnerable contact node between the susceptible ones.	Reduce the quantity of collected charge and thus reduce the probability of SEDNU to cause MBUs.	Utilizes more complicated intra-cell routing, therefore; increases fabrication cost.
Device resizing [60], High-density well contacts [116]	Gate	Increase the size of the critical nodes of a gate.	Reduce SER.	Increase fabrication cost, besides they do not guarantee reduce SET pulse width.
Reliable-aware logic synthesis [66]	Gate	Creating an improved cell library to replace cells with alternatives.	Reduce the SET width and the diffusion drain area of a circuit by 30% and 40%, respectively.	Impact system's speed performance higher than redundancy approaches.
DICE [74]	Circuit	Adding more transistors to protect the sensitive nodes in a latch or flip-flop.	Reduce the fabrication cost and design complexity, improves soft error roughly by 10-fold.	Vulnerable to DNUs under technology node scaling.
BISER [15]	Circuit	Use double conventional latches operating in parallel, C-element and weak keeper to mask soft errors and fight against leakage currents.	Improve fault masking resilience by 10-fold at 10.3% of power consumption penalty.	Consume high area overhead and still susceptible to MBUs.
SEM and STEM [85]	Circuit	Reduce the size of voting circuit by utilizing EDC circuit to detect errors and recover the system by rollbacking the correct data.	Average performance improvement of 26.58% compared to a conventional TMR and STEM outperforms SEM by 27.42%.	Incur area and power consumption overhead and increase the complexity of the design.
Temporal redundancy [2, 20, 117]	Circuit	Triple the sequential paths with simplex combinational logic path or repeat the operation and check for discrepancy.	Reduce the inherent overheads of spatial redundancy.	Incur the system's performance, depends on Eq. (2.3).
Spatial redundancy [110, 118, 119]	Module	Double or triple the logic paths in a system.	Masking both SEU and STE.	Incur more than 100% / 200% of area and power overheads.
Self-voting DMR [4]	Module	Double the logic datapaths and triple the sequential logic.	$10 - 24\%$ area improvement over the TMR.	Compromised speed performance and area/power overheads as compared to spatial/temporal redundancy approaches.
Diverse DMR (DDMR) and structural DMR [6, 87]	Module	Design diversity with pattern mismatch, utilizes two distinguished structures with equivalent functionality.	Consume 2.25x area compared to 3.003x of TMR implementation.	Require a discrimination circuit and identifies a pair of module implementations that produce pattern mismatch.
hCED [86]	Module	Use CED scheme and exploit the linear transform of the pipelineable applications.	Reduce power consumption by reducing the size of the checker circuit.	Cannot correct errors and needs condition to identify a block as fault-free.
EEC and parity [10, 48, 107]	System	Require extra bits of information be stored with the data in memory cells.	Incur less area, speed, and power overheads.	Vulnerable to MBUs, and ceasing the system clock for synchronization.
Combining ECC with interleaving scheme [114]	System	Separate data bits within each word adequately large.	Reduce the likelihood of MBU in SRAMs.	Complicated process and impacts area and power consumption.

From Table 2.5, it can be inferred that a system's performance is a tradeoff between the discussed issues, and selecting one of the discussed approaches to protect a design against soft errors depends on the application and the design constraints. In summary, as seen in Table 2.6, circuit-level mitigation approaches are the most dominant techniques to protect logic paths against soft errors since device-level techniques appear to become more difficult to implement and cost higher as device feature sizes are more shrunk. Meanwhile, effective alternative solutions that potentially offer intriguing features, such as being directly implementable on commercial technologies, able to tolerate SEU and MBU, while incurring acceptable area-energy and performance overheads under moderate cost are still sought to increase robustness of latching circuits at the circuit-level.

Table 2.6: Comparison between design-level abstraction of SE-mitigation approaches, where by each (✓, –) indicates relative (strength, limitation/weakness).

Approach	**Performance Degradation**	**Fault Masking Coverage**	**Complexity / Cost**	**Leverages Inherent Masking Mechanisms**
System-level	– low / medium	✓ high for both SEU and SET	– – moderate	– – – not exploited
Circuit/Module -level	– medium / acceptable	✓ high for both SEU and SET	– – moderate	✓ utilized considerably
Device-level	– – medium / high	– 100% for SEU not guaranteed for SET	– – – high	– – – not exploited

2.4 SER Recent Trends and Technology Challenges

Recent advancements in technology process, use of low-power technologies such as low-voltage operation, and the massive growth of device integration density have increased the susceptibility of integrated circuits to SER [16, 21, 26, 29]. Herein, some trends and challenges of recent CMOS device/circuit design constraints on SER are reviewed.

2.4.1 Trends of Technology Scaling on SER

In the literature, several research efforts [26, 27, 120, 121] have been carried out to investigate the trend of transient faults on the SER in CMOS logic designs under technology size scaling. Due to many divergent trends contributing to SET occurrences, the effect of technology scaling on the SET width is complicated [26]. Overall, an increase in SER is expected in current computing devices, despite decreasing the SEU rate per bit [10]. There exists several possible explanations to elucidate this issue. However, the most distinctly possible explanations are that under CMOS scaling, VLSI chips have become denser and smaller, therefore, its immunity towards particle strikes decreases significantly. With the massive growth in integration capacity in advanced CMOS technology, the susceptible nodes of the drain area have been increased, thereby leading to higher SER [10]. Additionally, in the earlier technologies, the main contributor of soft errors was indirect ionization of neutrons. Even though the neutrons by themselves are unable to cause direct ionization, however, the high energy neutron collision that interacts with semiconductor materials induce sufficient ionizing energy depositions, which can cause SEUs. Therefore, the secondary particles of high energy neutrons have the largest effect on Q_{crit} at sea level. However, as the device feature size is shrunk, Q_{crit} becomes smaller, alpha particles contribute increased SER in CMOS SRAMs, while secondary particles of neutrons still dominate SER in latching circuits [39].

Overall, the most distinctly possible explanations that elucidate why the current trend of soft-error towards higher rates is that under technology scaling, the number of transistors per unit area doubles, thus SER per unit area continually increases even though SEU per bit has decreased [10, 17, 26]. Thus, soft errors have been pronounced in the literature as predominant sources of reliability degradation due to VLSI chips' immunity towards particle strikes decreases significantly. Nevertheless, the soft-error historical studies in [26] and [16] report that the only fact is that SET rate has not yet exceeded the SEUs failure rate of a storage element. However,

single-event induced multiple node upset have become more pronounced due to reducing device dimensions. Therefore, the designers of microprocessors and memory structures should take into account MCUs/MBUs in the SER protection techniques.

2.4.2 Trends of Voltage Scaling on SER

Supply voltage (V_{DD}) scaling is widely recognized as the most effective strategy to reduce power dissipation/consumption for VLSI logic devices. In fact, Near Threshold Voltage (NTV) operation slightly reduces supply voltage to 100-200mV above the threshold voltage of the transistors, allowing for significant improvement in energy-efficiency [122], significant energy savings as compared to operating at nominal voltage [123]. However, the critical charge needed to cause a transient fault decreases linearly as V_{DD} is scaled, and the SER exhibits an exponential effect based on critical charge [10, 50]. Such trends are consistent with decreasing feature sizes due to technology scaling [124].

For instance, [48] predicts that the SER of logic circuits per die will become comparable to the SER for unprotected memory elements, which was later verified through experimental data measured from a microprocessor [10]. Operation at NTV is predicted to aggravate these trends. For example, [10] reports that SER increases by roughly 30% per each $0.1V$ decrease as V_{DD} is decreased from 1.25V to 0.5V. Moreover, both simulation results and experimental results at the 40nm and 28nm technology nodes concur that SER increases by two orders of magnitude when V_{DD} is reduced from 0.7V to 0.5V [10]. Thus, microprocessor energy reduction techniques, such as Dynamic Voltage Frequency Scaling (DVFS), might suffer higher SER, i.e., an increased rate of both SET and SEU, due to reducing V_{DD} to minimize energy consumption, and therefore, negatively impact the microprocessor's reliability [10].

2.4.3 Trends of Multi-Gate FinFET Devices on SER

The non-planar structure devices, i.e., fin-typed Field-Effect-Transistors (FinFETs), have been introduced to continue scaling and also show improved low-leakage and high performance over planar MOSFET devices [30, 31]. However, they also can offer alternatives that achieve increased tolerance to transient faults, i.e., SET and SEU [16]. The benefits of non-planar multi-gate devices on reducing SER in CMOS circuits have been emphasized in numerous studies [30–34, 43–45, 125, 126]. In [44], it was reported that FinFET devices have been introduced as alternatives with a 50% reduction of threshold voltage fluctuation of 16nm-gate HK-MG bulk FinFETs compared with the results of planar MOSFETs. Additionally, even though spatial MBU feasibility has increased due to technology scaling, the experimental results in [34] showed that non-planar devices offer a means to reduce SER in logic paths and MCU in memory arrays. For example, 22nm tri-gate technology is shown to reduce neutron-induced SER at nominal voltage from 1.5-folds to 4-folds and alpha particle SER in excess of 10-folds compared to a 32nm planar process. Moreover, Fang and Oates [125] projected that the neutron-induced SER of bulk FinFET SRAM cells is roughly 6% compared to the SRAM that is configured with non-planar devices. In contrast, even though scaling the transistor volume helps in reducing the susceptible area, there exists an overall agreement that SER per chip increases with nanoscale device technology generations [10, 14, 16, 21]. This is due to lower Q_{crit} and higher device density per unit area, i.e., roughly doubling, which leverages a higher strike probability. Consequently, low energy particles can deposit an adequate amount of charge, i.e., sufficient Linear Energy Transfer (LET), to cause soft errors.

2.5 Summary

The revolution in chip manufacturing has placed multicore and many core microprocessors at the forefront of delivering high-performance and energy-efficient nano-electronic devices to servers

used in cloud and High-Performance Computing (HPC). These systems employ billions of transistors within large-scale VLSI chips to elevate performance. However, these advancements have also continually augmented the impact of Single-Event Transient (SET) and Single-Event Upset (SEU) occurrences which precipitate a range of Soft-Error (SE) dependability issues. In this chapter, a wide range of soft error resilience mitigation techniques that exploit specific attributes of different design levels for energy-saving under a reliable protection levels are discussed. These techniques have been adopted to mitigate soft errors across different levels of design abstraction. Determining which radiation-hardening scheme achieves better performance is a tradeoff between the design requirements and their challenges. However, SER mitigation techniques at the circuit/module-level can be convenient since not all SETs will cause an upset in logic paths. Overall, all the reviewed SE-mitigation techniques involve high area-power overhead and/or circuit performance degradation. Furthermore, as technology trends for increased transistor density continue, SEU-immune schemes have become inadequate for robust designs. Thus, more powerful SER mitigation schemes, and means for cost-efficient and MBU/DNU-tolerance, that consider layout modifications and/or circuit design topology with reduced area-power overhead and circuit performance degradation are sought. Considering the complexity of large-scale mission-critical systems, robustness and energy-efficiency are the major challenges that need careful investigation for tradeoff. Developing SE-resilient latches to budget or compensate for these transient faults is essential for resilient systems. Addressing these reliability concerns is paramount to successful deployment of future reliable circuits and systems with deeply-scaled devices operating at low-voltages.

CHAPTER 3: TEMPORAL SELF-VOTING CHECKERS: ENERGY VS. RESILIENCE TRADEOFFS

In this chapter, we developed two circuit-level techniques, namely Temporal Self-Voting Logic (TSVL) and Hybrid Spatial and Temporal Redundancy Double-Error Correction (HSTR-DEC), to prevent the effects of soft errors in logic circuits, occurring due to Single-Event Upset (SEU) or Single-Event Transient (SET). TSVL and HSTR-DEC circuits can be utilized to improve the reliability of a logic path with minimal impact on circuit delay while achieving a high coverage and cost-effective SEU handling as compared to the previously presented redundancy-based soft-error masking approaches.[1]

3.1 Energy-Efficiency vs Fault Resilience

SE-mitigation techniques are utilized for designing dependable systems so that high reliability and availability are realized while directly countering desirable attributes such as, high speed, low power, and minimal area overhead. However, adding extra logic to harden a design against soft error effects increases the number of susceptible nodes, leading towards the reduction of the system's overall performance. Hence, the challenge is not how to realize a high reliability, however, achieving it with minimal area and energy overheads and speed degradation is the aim. Considering all these factors, tradeoffs between fault resilience and energy consumption were carried out in designing the proposed soft error handling approaches, so that slightly sacrificing fault making coverage reduces the energy consumption of proposed schemes. The primary contribution of these SEU-tolerant approaches is that they eliminate error masking from the combinational datapath logic, thus, area and energy overheads are significantly reduced. It is worth noticing that the

[1] © 2016 IEEE. Reprinted, with permission, from [67].

presented approaches achieve resilient fault masking coverage, higher than 96% for the worst case.

3.2 The Proposed Approaches

Achieving high reliability against transient faults poses significant challenges due to the trends of technology and voltage scaling. Thus, numerous soft error mitigation techniques [2, 4, 85, 86] have been proposed for masking Soft Error Rate (SER) in logic circuits. However, most soft error suppression approaches have significant overheads in terms of area, power consumption, and speed performance degradation. Herein, we introduce two novel efficient and cost-effective techniques to harden the logic paths against soft error effects. The proposed approaches concentrate on protecting the sequential logic elements, i.e., flip-flops or latches, from an upset while achieving significant area and power saving.

3.2.1 Temporal Self-Voting Logic (TSVL) Approach

In conventional Dual Module Redundancy (DMR) approach, the outputs of two identical modules are compared, and an error is detected where there is a discrepancy between them. Conventional DMR scheme is used to detect when a soft error occurs, without error correction, since voting cannot determine which of the two modules is error-free [78]. However, the proposed scheme, TSVL, can be utilized to detect and correct any upset in a flip-flop while incurring acceptable area, power, and performance penalties. As seen in Figure 3.1, the design is able to detect and correct any SEU that hits one of the flip-flops or an SET that generates, propagates, and eventually is captured by one of the flip-flops. Herein, there are three possible scenarios in which an upset will occur. The first scenario is when an SET hits the combinational logic preceding the flip-flops and then propagates and arrives at the setup/hold time of the first/original flip-flop, overlapping

with window of vulnerability. This will upset the original flip-flop, while the redundant flip-flop is error-free as long as it is triggered by a clock delayed by a phase shift greater than the width of the transient pulse. Thus, the comparator (XOR-gate) will assert the $Error_{SEU}$ signal to indicate the occurrence of an upset, making the self-voter circuit determine the final output.

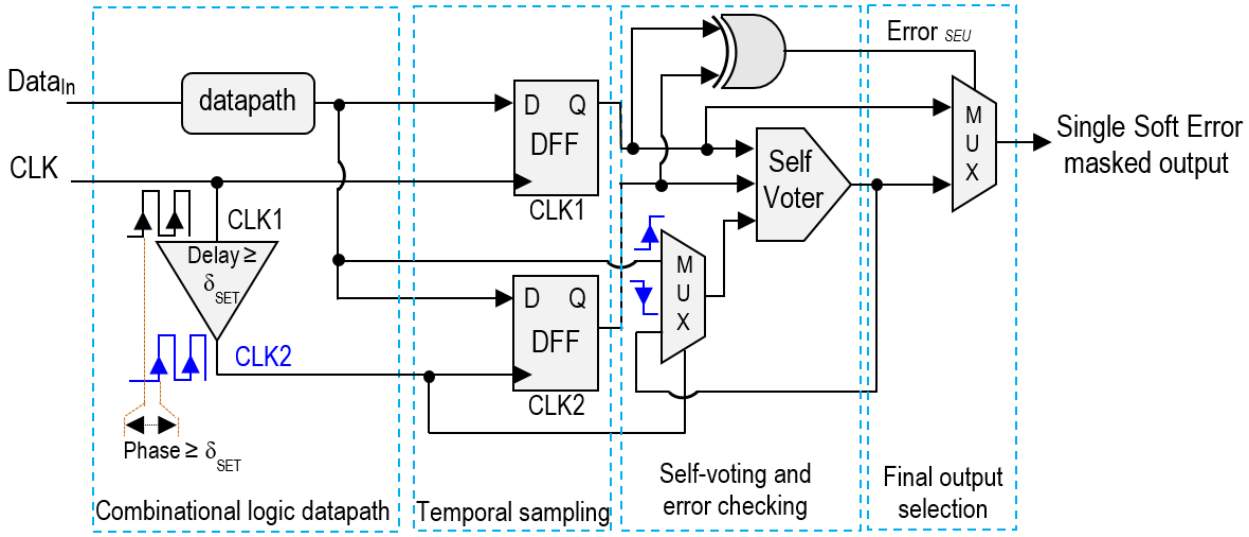

Figure 3.1: Temporal Self-Voting Logic (TSVL) approach.

Hence, the third input of the self-voter is fed from a MUX that receives its first input directly from the datapath and the second input from the feedback of self-voter circuit; an SEU will be masked by the self-voter circuit regardless of whether it occurs in the original or in the redundant flip-flop. This is because the third input of the self-voter determines the output when the external two inputs mismatch. Since the third input of self-voter circuit relies on the output of the first MUX that drives its output based on the delayed clock rate (CLK2), the self-voter's third input will be switched alternately between the direct combinaltionl datapath, with the rising edge, and the feedback of self-voter circuit, with the falling edge. Thus, when an SET hits the combinational logic and is then captured by one of the flip-flops, its duration is significant. If it will have passed by the time that the first input's rising edge arrives to the first MUX, then it will be ignored as long as the delay of CLK2 is larger than the generated SET. Therefore, the correct data will pass

as the self-voter's third input with the rising edge. The purpose for using a MUX circuit at the input of self-voter is to update the previous state of the self-voter circuit with the rising edge of the delayed clock (CLK2). Figure 3.2 depicts the validity of the proposed scheme, as two individual upsets occur in the original and redundant flip-flop, get corrected, and legitimate data is passed to the final output.

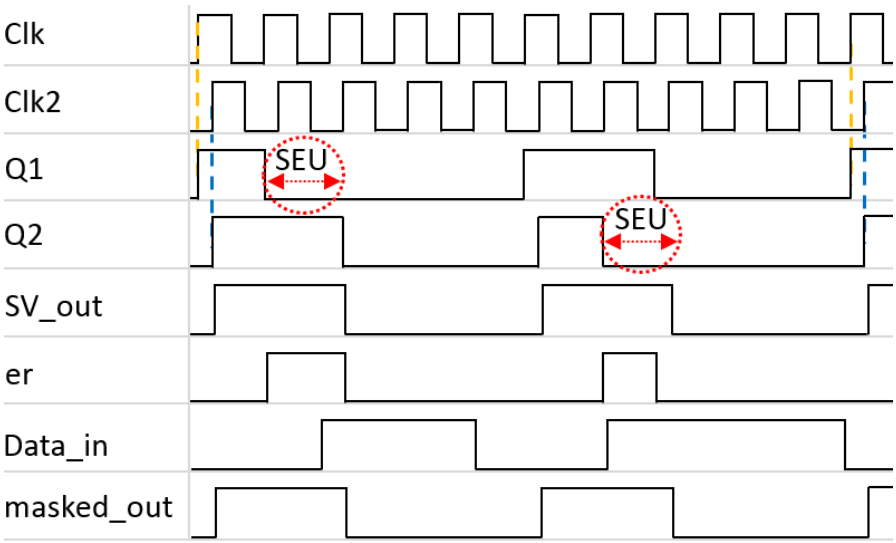

Figure 3.2: Timing diagram of TSVL approach with SEU recovery occurred in the original, $Q1$, and redundant, $Q2$, flip-flop consecutively, `masked_out` is an inverse of `data_in`.

As a result, irrespective of whether the upset occurs in the original or redundant flip-flop, the final output is always correct as long as the third input of self-voter circuit alters based on CLK2. The second scenario is when an SET hits inside a flip-flop, an error (SEU) will be observed, and it will be masked in the same way as stated above. The last scenario is when the first MUX or the XOR-gate is hit by a transient pulse and in this case, the final output is not effected as long as the output of both flip-flops are correct; therefore, any one of them can be selected as the final output. However, an SET that hits the second MUX circuit, that determines the final output, will cause a momentary glitch in the final output, as in the majority voter of a TMR. Note that signal CLK2 is

generated by buffering the main clock to add phase delay greater than the SET pulse width, which is a popular practice in the design of reliable circuits. The main difference between the proposed approach and the Self-Voting DMR approach presented in [4] is that TSVL functions are based on a simplex combinational datapath and double flip-flops, whereas Self-Voting DMR uses DMR for the combinational logic portion and TMR for the sequential logic portion, thus TSVL achieves significant reduction in area and power penalties as compared to the Self-Voting DMR approach; this will be discussed in Section 3.3.

As compared to spatial and temporal redundancy approach flip-flop based designs, TSVL provides a performance compromise between TMR and temporal redundancy approaches. The proposed scheme (TSVL) consumes less area (by two redundant logic modules) than TMR and provides higher performance in term of speed than temporal redundancy (roughly half the speed degradation of the temporal redundancy scheme). Overall, the delay of TSVL approach is given by:

$$\delta_{TSVL} = \delta_{critical} + \delta_{voter} + 2 * \delta_{Mux} + \delta_{SET} \tag{3.1}$$

$\delta_{critical}$, δ_{voter}, and δ_{SET} were discussed in chapter 2.

Based on Eq. 2.5, in chapter 2, the area overhead for TSVL approach is given by:

$$A_{TSVL} = A_{seq} + A_{voter} + 2 * A_{Mux} \tag{3.2}$$

Equation 3.1 considers the longest datapath delay. Thus, to meet the timing constraints of a radiation hardening technique, its main clock period must be greater than, or equal to, the longest logic datapath delay.

3.2.2 Hybrid Spatial and Temporal Redundancy Double-Error Correction Approach

Recently, CMOS technologies have been deeply-scaled, thereby a single particle strike that hits a sensitive node can affect more than a single storage element at a given instance of time [95]. This implies that the probability of double or even multiple errors occurring simultaneously has become high. Thus, soft error in flip-flops has become a concern due to the effect of charge sharing between the adjacent nodes [95]. Additionally, circuit-level mitigation techniques introduce reduced benefits under technology scaling. For instance, in [115], it is predicted that, as the spacing between the sensitive nodes is reduced, the vulnerability of hardened flip-flops increases. Therefore, optimized solutions which offer multiple error correction while minimizing performance degradation and energy overhead, imposed by the extra logic for protection, are sought.

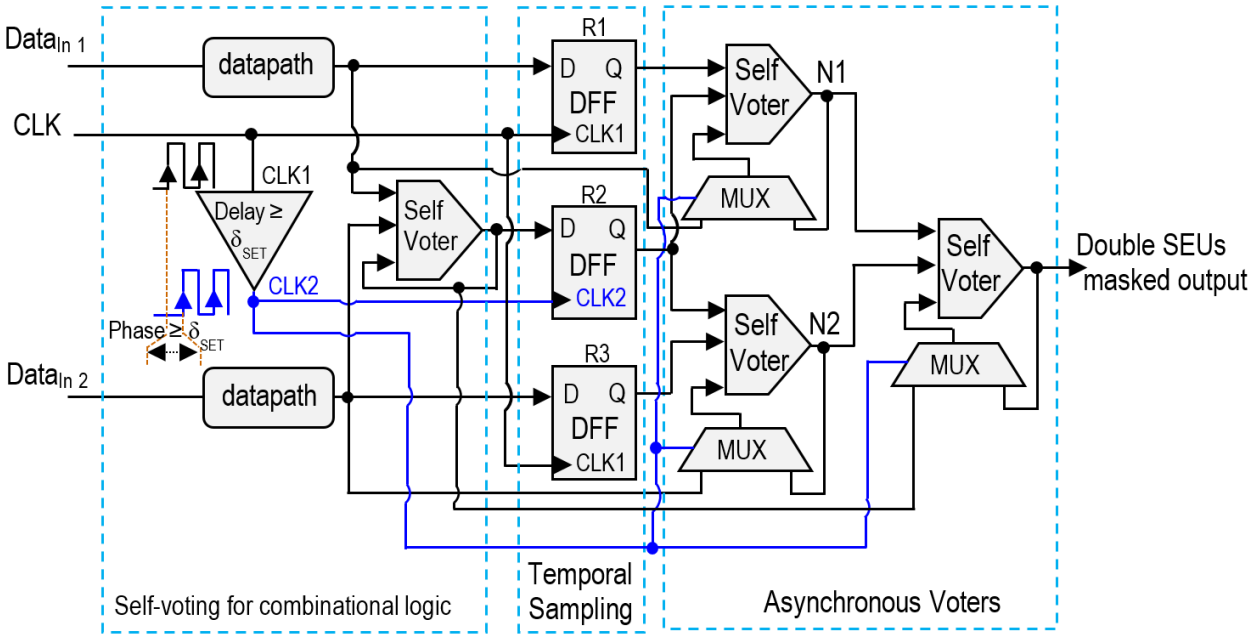

Figure 3.3: Hybrid Spatial and Temporal Redundancy Double-Error Correction approach.

Based on the issue stated above, we develop a new technique, namely Hybrid Spatial and Tempo-

ral Redundancy Double-Error Correction (HSTR-DEC), for soft error tolerance. As illustrated in Figure 3.3, the proposed approach utilizes DMR circuit for the combinational datapath with TMR for the sequential portion, i.e., flip-flops. The final output is determined by voting based on two masked outputs, nodes N1 and N2, and the masked output from the original and redundant datapaths. The proposed scheme significantly improves the error resilience as compared to previous traditional redundancy approaches due to its ability to tolerate double SEUs simultaneously, i.e., Multi-bit Upsets (MBUs) are tolerated. Moreover, HSTR-DEC approach can surpass not only the SEUs, but also can mask SETs that generate through datapath logic and are eventually captured by one of the flip-flops. Our approach has been motivated by the original technique proposed in [4], and the proposed changes have been designed by evaluating the limitations of error resilience, i.e., Multi-Bit Upsets (MBUs), regarding area and delay overheads. Table 3.1 shows the possible scenarios of soft error (SEU) that HSTR-DEC approach can tolerate. In case a single SEU occurs, the proposed approach is able to mask an upset regardless of which register the error occurs.

Table 3.1: Possible error scenarios of HSTR-DEC approach; (NE = No Error; UE = Upset Error).

♯ of SEUs	Case	R1	R2	R3	N1	N2	Masked Output
Single	I	UE	NE	NE	✓	✓	✓
	II	NE	UE	NE	✓	✓	✓
	III	NE	NE	UE	✓	✓	✓
Double	I	UE	UE	NE	UE	✓	✓
	II	UE	NE	UE	✓	✓	✓
	III	NE	UE	UE	✓	UE	✓

On the other hand, in case there are two SETs generated and propagated through the datapath at the same time, they might be captured by R1 and R3 only, whereas R2 will capture the legitimate data since it is triggered by a clock which is delayed by a phase shift greater than the generated transient pulse. There exist three possible scenarios for two SEUs to occur simultaneously. In case I, both register R1 and R2 are upset, therefore, the output of the first self-voter circuit, node N1, is incorrect while node N2 is correct due to the second self-voter votes based on R2, R3, and the

redundant direct datapath during the rising edge of the clock. Thus, N2 is error-free as long as both R3 and the redundant datapath are correct. The same scenario, i.e., case III, occurs when both R2 and R3 get upset. In case II, register R1 and R3 get upset, meanwhile nodes N1 and N2 are correct. This is because both self-voter circuits vote based on R2, which is error-free, and the original and redundant datapaths which are also error-free as long as the width of SET is less than 200 psec. However, the proposed scheme is unable to recover when all the flip-flops, i.e., R1, R2, and R3, become upset at the same time. This indicates that the proposed scheme fails to recover only when the energetic particles strike inside the flip-flops and are adequately large to flip the bit state at all registers during the same clock period.

3.3 Experimental Results and Analysis

To assess alternatives that can mask Soft Error Effects (SEE) under the impact of transistor scaling, Synopsys Design Compiler was used to synthesize the proposed approach using 15nm technology process based on NanGate open source library. Since TSVL does not require any modifications or adoptions at the device-level, the standard cell libraries were used to implement the circuits. Thus, no complexity or additional fabrication cost will be associated with the proposed scheme under technology generations. Additionally, area and energy costs are minimized in the generation of the delayed clock signal (CLK2) by using buffers that introduce maximum delay with minimum power consumption and area usage.

3.3.1 Quantifying Fault Resilience for Proposed Techniques

The ability of the proposed hardening approaches to suppress soft errors, was tested by randomly injecting transient and upset faults at the gate-level design. As a case study, a set of ISCAS89

benchmark circuits, with a different combinational and sequential logic ratio, have been synthesized using Synopsys Design Compiler and validated with fault injection technique, presented in [127], where 7000 faults, both SET and SEU, were injected into the gate-level Verilog code describing the benchmark circuit constructed with the proposed approach. SEUs were injected into the input node of a flip-flop with the rising edge, whereas SETs were injected into a node of a logic gate, and they were injected at arbitrarily chosen locations and times. The fault coverage of the proposed approach is listed in Table 3.2 for a set of ISCAS89 benchmark circuits constructed with TSVL using 15nm technology process.

Table 3.2: Fault coverage analysis of TSVL approach.

Benckmark Circuit	♯ of Cells	♯ of Injected Faults	♯ of Masked Faults	♯ of Unmasked Faults	Fault Coverage %	SV-DMR [4] FCPR	TSVL FCPR
$S27$	13	500	487	13	97.4	20.40	23.75
$S349$	127	4000	3883	117	97.075	3.51	4.35
$S444$	163	7000	6814	186	97.34	2.13	2.5
$S838$	418	7000	6798	202	97.11	1.73	2.21
$S1423$	683	7000	6779	221	96.84	0.68	0.804
$S9234$	1369	7000	6737	263	96.24	0.365	0.455

The analysis of simulation results indicates that TSVL is able to recover from soft errors by 100% of SEUs and 97% of SETs, thus, on average, it can roughly suppress 99% of soft errors. In addition, since SER mitigation techniques offer a tradeoff between the error resilience and overheads of protection techniques, we calculate a new metric called Fault Coverage Power Ratio (FCPR) by dividing the fault coverage by the power consumption, as the goal is to maximize the fault coverage and minimize the power consumption. Thus, the higher the value of FCPR, the better the design. As seen in the last two columns of Table 3.2, TSVL realizes higher FCPR for all the selected benchmark circuits, regardless whether the combinational or the sequential logic ratio is higher. Also, TSVL achieves higher FCPR for the small circuits. Notice that the unmasked faults in the $S1423$ and $S9234$ are large because these benchmark circuits utilize large number of flip-flops, 74

and 228 respectively, thus the probability of an SET hitting the second Mux circuit is increased. Thus, the proposed approach can be advantageous for soft error tolerant designs constrained with a tight energy consumption budget with a minimal impact on the reliability by 1% for the soft error effects.

3.3.2 Performance Evaluation for Proposed Techniques

In this section, area overhead, power consumption, and speed degradation are quantified for the proposed techniques. The area overhead for the TSVL is estimated by measuring the implementation area in um^2, and it turned out that the proposed scheme requires 54.53 um^2, whereas 66.5, 64.9, and 43.4 um^2 is needed to configure $S27$ benckmark circuit with TMR, Self-Voting DMR, and temporal redundancy approach, respectively. Figure 3.4 shows the area overhead for a set of ISCAS89 benchmark circuits with a distinct combinational and sequential logic ratio. It can be seen that the proposed approach imposes an area overhead comparable to the temporal redundancy, while achieving significant reduction as compared to Self-Voting DMR or TMR approach. In addition, Self-Voting DMR approach realizes an area reduction when the sequential logic portion is high, while the proposed approach saves significant area either if the sequential logic portion is high or if the combinational logic portion is high. On average, TSVL approach imposes less area overhead by 22.02% and 36.84% as compared to Self-Voting DMR approach, presented in [4], and TMR, respectively. In terms of power consumption, TSVL realizes power consumption improvement of 20.1% and 35.55% over the Self-Voting DMR and conventional TMR approach, respectively, as depicted in Figure 3.5. Therefore, for applications which seek to protect a design against soft errors with constant energy budget, it is advantageous to utilize TSVL approach as it imposes high level of protection, besides achieving significant reduction in area-energy overhead as compared to other hardening redundancy-based techniques.

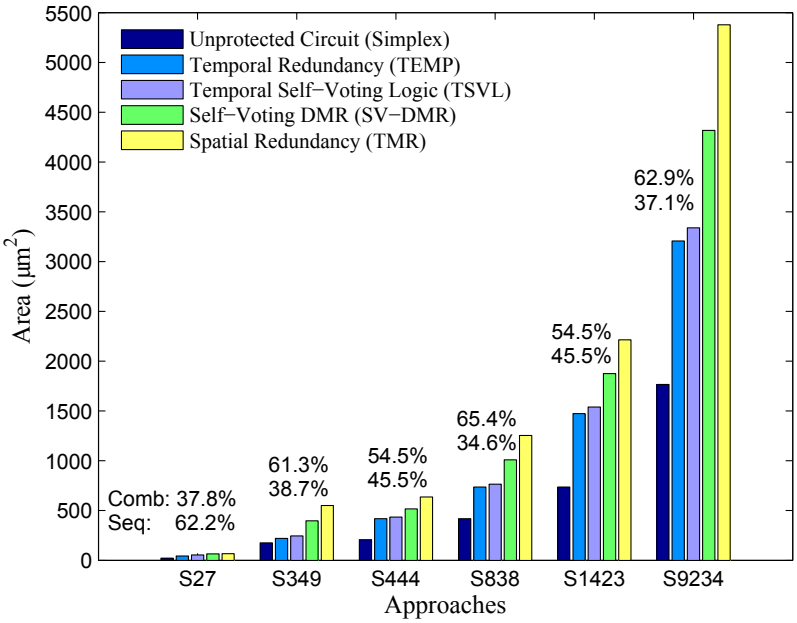

Figure 3.4: Area overhead evaluation of redundancy-based soft error masking techniques.

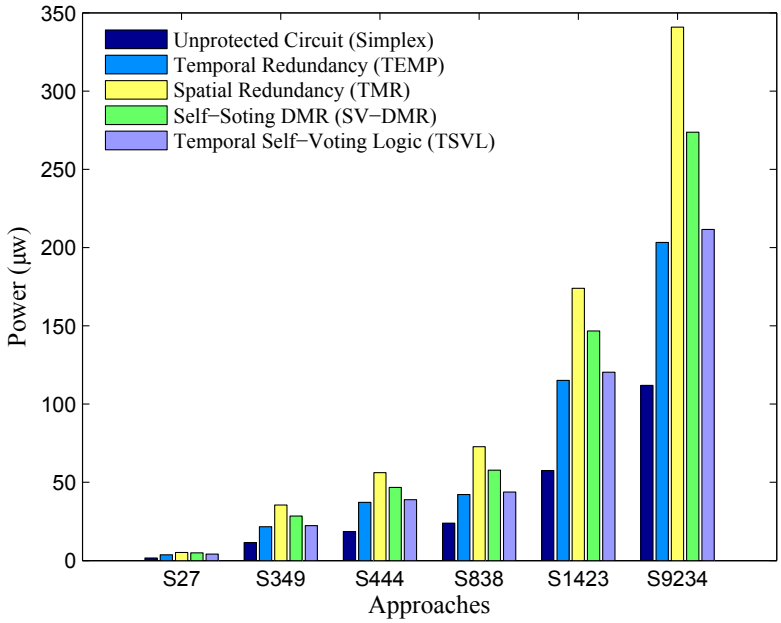

Figure 3.5: Power consumption analysis of redundancy-based soft error masking techniques.

In terms of performance, TSVL provides a comparable speed performance to that of Self-Voting DMR approach, shown in Figure 3.6, as both incur a phase delay within a period that depends on the SET pulse width. Overall, TSVL outperforms Self-Voting DMR by 2.15%. A transient pulse of 200 psec was considered in order to generate CLK2 that triggers the second and third flip-flop of TSVL and Self-Voting DMR approach, respectively. Thus, an SET pulse within a width larger than the considered phase delay might be captured by a flip-flop if it arrives at the setup/hold time thereby causing an upset in that storage element, unless it is masked via the electrical or logical masking mechanism.

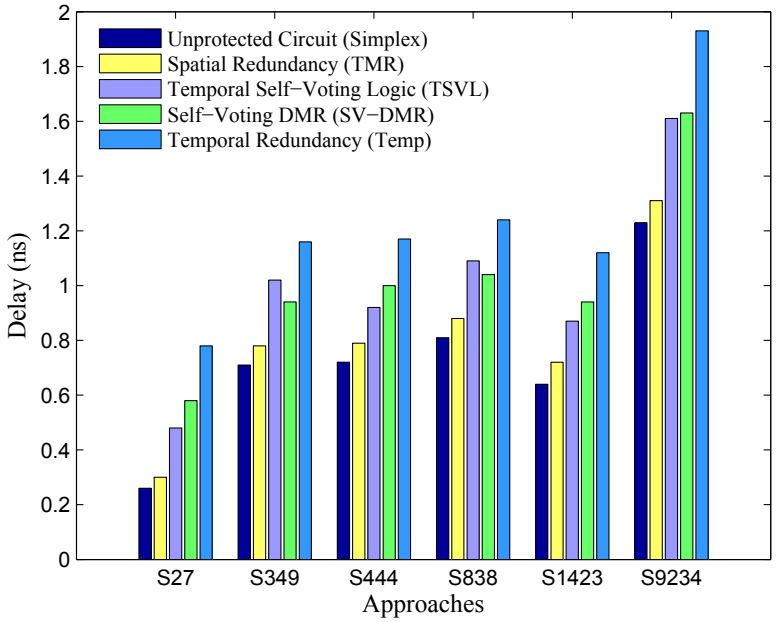

Figure 3.6: Performance evaluation of redundancy-based soft error masking techniques.

Interestingly, HSTR-DEC approach is able to tolerate an SET within a width wider than 200 psec. However, in this case the transient pulse should occur in the original datapath only, $Data_{in1}$, and the redundant datapath should be error-free at the same time, and vice versa. This indicates that HSTR-DEC is able to tolerate a single SET with a width larger than the considered phase delay

as long as the other datapath is correct. On the other hand, HSTR-DEC scheme is able to tolerate two simultaneous SETs within a width of up to 200 psec. While TMR, temporal redundancy, and self-voting DMR approaches achieve complete fault masking coverage for SEU in single module simultaneously, HSTR-DEC approach realizes complete fault masking coverage for SEU in single or double module concurrently. In addition to its increased level of reliability, the proposed scheme consumes acceptable extra logic for error masking. The evaluation results indicate that HSTR-DRC approach reduces area and power overheads roughly by 18.2% and 16.83%, respectively, and imposes the performance by 20.17% as compared to conventional TMR using 15nm technology.

Table 3.3: Mean area-power consumption and delay of redundancy-based approaches w.r.t. simplex design.

Design Implementation	**Normalized Area**	**Normalized Power**	**Normalized Delay**
Spatial Redundancy	$3.14x$	$3.11x$	$1.19x$
Temporal Redundancy [2]	$1.88x$	$1.83x$	$1.69x$
Self-Voting DMR [4]	$2.48x$	$2.46x$	$1.4x$
Proposed Approach (TSVL)	$1.96x$	$1.91x$	$1.37x$
Proposed Approach (HSTR-DEC)	$2.6x$	$2.54x$	$1.43x$

On the other hand, it improves speed performance by 15.38% while incurring 19.23% of area overhead as compared to temporal redundancy approach. Likewise, it achieves comparable overheads in terms of power, area, and delay as compared to previous works of hybrid redundancy approach, presented in [4], as both utilize duplicated datpaths and triplicated flip-flops. However, our approach is able to tolerate double upsets that occur at the same time, whereas conventional TMR, temporal redundancy, and SV-DMR are unable to detect and correct double upsets simultaneously. Table 4.2 lists the area, power consumption, and speed degradation penalties of the redundancy-based soft error mitigation techniques, normalized to the simplex (non-redundant) design. In addition to increased levels of fault coverage, the proposed approaches are considered a compromise solution ranging between the spatial and temporal redundancy approaches, as they

realize a high protection against soft error effects, MBUs tolerance for HSTR-DEC, and impose acceptable overheads in terms of area, power, and performance degradation.

3.4 Summary

As technology continues to scale, CMOS logic becomes denser, thus its immunity towards particle strikes decreases significantly. In this work, efficient and cost-effective redundancy-based Soft Error (SE) handling approaches are proposed, namely TSVL and HSTR-DEC. They are utilized to prevent the effects of soft errors in logic paths, occurring due to SEU or SET. The proposed approaches are validated by injecting transient and upset glitches. Experimental results indicate that TSVL approach can cover soft errors, on average, by roughly 99% while realizing an improvement of 22.02% and 2.15% for area and speed degradation compared to the previous Self-Voting DMR approach. Meanwhile, HSTR-DEC approach realizes complete masking coverage for single and double SEUs while incurring comparable area and delay overheads as compared to the prior hybrid redundancy approach, aside from its energy-efficiency as compared to conventional TMR approach.

CHAPTER 4: IMPACT OF PROCESS VARIATION IN THE NVT REGION ON SOFT ERROR RATE

In this chapter, we investigate the effect of both technology size scaling and delay variation at Near-Threshold Voltage (NTV) region on redundant systems, regarding energy versus performance tradeoffs within an iso-energy constraints. We develop new results for the evaluation of alternatives to mask Single Event Transients (SET) in combinational logic and Single Event Upsets (SEU) in storage elements for three commonly utilized redundancy approaches, namely, spatial, temporal and hybrid of both spatial and temporal. The performance, and energy impact of each approach is quantified at NTV operation. Additionally, the impact of increased effect of threshold voltage variation at NTV is assessed for all redundant systems. We also investigate the effect of technology scaling by comparing the energy and performance variation of 45-nm MOSFET planar and 16-nm High-κ/Metal Gate (HK-MG) bulk FinFETs structures as modeled by PTM NanGate open source library via simulations in HSPICE. Monte-Carlo simulations reveal the influence of the threshold voltage (V_{th}) variation in redundant systems at NTV to mask a transient pulse in combinational logic and/or an upset glitch in a storage element. The primary insight from this work is identification of redundancy-based hardening techniques that can deliver increased benefits in terms of the *Fault Coverage Energy Ratio (FCER)* for the leveraged tradeoffs within iso-energy constraints at Near-Threshold Voltage (NTV). We demonstrate how the FCER metric can be used as an optimization parameter to guide circuit synthesis to meet performance and robustness goals. Finally, the impact of design diversity on spatial and hybrid redundancy at NTV is assessed in terms of FCER and delay variation to form overall recommendations regarding soft-error mitigation at NTV.[1][2]

[1] © 2017 IEEE. Part of this chapter is reprinted, with permission, from [29].

[2] © 2018 IEEE. Part of this chapter is reprinted, with permission, from [69].

4.1 Energy-Efficiency Through Near-Threshold Computing

Supply voltage (V_{DD}) scaling is widely recognized as an effective lever to reduce the power dissipation and energy consumption for VLSI logic devices. Total energy consumption of CMOS logic circuits is determined by dynamic and static energy components, which are both dependent on supply voltage. The dynamic energy component has a quadratic relationship with V_{DD}, whereas the static components have a linear dependence. This property can be exploited in a harsh environment, where some remote accessed systems might operate with serious constraints or power consumption budgets. For instance, low power designs can be protected more adequately by employing the redundancy approaches at lower V_{DD} and lower performance. Operating at low-voltage such as below the transistor's threshold voltage V_{th}, which is an order of magnitude larger than thermal voltage, can result in highly undesirable exponential increase in delay. The leakage energy also begins to dominate at a certain point, such that designs within this *sub-threshold region* have limited applicability. Hence, operation in the *near-threshold region* is sought, as it provides an energy-efficient operating tradeoff against delay. In Near-Threshold Voltage (NTV) region, the V_{DD} is set to be slightly above the V_{th} (100-200mV) [122] of the transistors to provide a 10X delay improvement compared to operation in the sub-threshold region with only a 50% reduction in energy savings [123]. Taking all of these factors into consideration, NTV can be preferred to provide up to 6X energy savings with reasonable performance overheads as compared to operating at nominal voltage [123], [128]. Even though NTV computing offers an attractive approach to balance power consumption versus delay for applications within iso-energy constraints, such as for high-performance computing [123], it can result in reliability implications as the fault masking coverage is reduced. In the literature, several studies have reported the impact of voltage scaling of integrated circuits to soft errors where millions of transistors are used to construct the ICs. Herein, we assess the relative performance of temporal and spatial alternatives from 800mV to 500mV using a 16-nm Predictive Technology Model for Multi-Gate transistor (PTM-MG) li-

brary.As near-threshold computing becomes mainstream, more mission-critical applications will seek to embrace this mode of operation for the energy benefits with maximum reliability. The potentials and pitfalls of near-threshold computing for mission-critical applications, such as on-line banking systems, avionics operation and control systems that can have a severe impact if they fail [101], are end-product impacts of this work.

4.2 Trends of SER in NTV Region

Operation with V_{DD} in the NTV region is sought for highly-scaled CMOS logic circuits due to its balancing of minor performance degradation relative to its significant power savings [129]. While NTV offers an attractive approach to balance energy consumption versus delay for power-constrained applications such as high-performance computing, there is a need to evaluate its reliability implications through increase in soft errors [122] and performance variation due to higher impact of threshold voltage variation [130]. The SER is exacerbated by two complex interacting factors: technology scaling and reduced V_{DD}. In particular, radiation-induced SEUs which cause soft errors can increase significantly in this operating region [10]. These soft errors are expected to increase at NTV and will be exaggerated with manufacturing-induced Process Variations (PV) due to technology scaling [10, 129]. Additionally, reducing V_{DD} will increase the impact of SETs since the amount of critical charge (Q_{crit}), which needs to be collected in order to change the state of an output node at a logic cell, is reduced. Primarily, the critical charge needed to cause a failure decreases as V_{DD} is scaled and SER has an exponential dependence on Q_{crit} [10]. As a result, soft errors during NTV operation can affect memory elements and cause arbitrary bit flip(s) in the stored data (SEU), or transient errors can propagate through a logic path and eventually become latched by a latch/flip-flop circuit (SET) to cause erroneous computations. Thus, the trend of SER is expected to increase at NTV operation of deeply-scaled VLSI systems.

4.3 Technology Scaling and PV Trends for SET

In the literature, some predictions of the transient pulse width trends under technology process scaling have been conflicting [9]. In fact, the impact of technology scaling on SETs is found to be complicated by several divergent trends contributing to SET characteristics, such as reducing Q_{crit}, Propagation-Induced Pulse Broadening (PIPB), pulse quenching, n-well contact area and spacing (to reduce parasitic bipolar effect), trends of datapath masking mechanisms, charge sharing, and source of radiation/particles [9, 26, 27]. However, the results presented in [9] assist in elucidating some of the discrepancies in SET pulse widths measurements presented by various researchers for the last decade. Gadlage *et al.* [9] experimentally measured the transient pulse widths in a bulk 130, 90, and 65-nm CMOS technology, and they reported that SET pulse width overall was found to be reduced under technology scaling for the conditions which they evaluated. However, an overall increase in SER is expected in recent computing devices, even though the SEU per bit has decreased due to lower Q_{crit} and higher device density per unit area, i.e., roughly doubling, which results in a higher strike probability, even in deeply-scaled non-planar devices that exhibit increased tolerance to particle strikes. Moreover, the analyses of the previous soft error studies have predicted that soft error in combinational logic and latches/flip-flops would dominate the overall CMOS chip errors, and it will exacerbate under technology process scaling [16]. Particularly, in [15], it is predicted that SER contribution of logic circuits to total chip SER, roughly estimated at 60%, exceeds SRAM which contributes 40% of SER during execution. Therefore, SER logic-based should be emphasized for effective SER mitigation [27].

On the other hand, to alleviate the Process Variation (PV) associated with technology scaling, FinFET devices have been introduced as an alternative with 50% reduction of threshold voltage fluctuation for 16-nm-gate HK-MG bulk FinFETs compared with the results of the planar MOSFETs [44]. In particular, performance improvement of nano scale CMOS devices requires not only

eliminating a variety of fabrication challenges (systematic or extrinsic variation) but also mitigating random or intrinsic variation effects including Random Dopant Fluctuation (RDF), Line-Edge Roughness (LER), Interface Traps (ITs), and Work-Function (WK) variations, which are crucial for device characterization of nanometer scale CMOS planar and FinFET structures [31, 43]. The non-planar devices offer a means to reduce SER. For example, 22-nm tri-gate technology is shown to reduce neutron induced SER at nominal voltage from 1.5-folds to 4-folds and alpha-particle SER in excess of 10-folds compared to a 32-nm planar process [34].

In this work, intra-die variations for both CMOS 45-nm planar and tri-gate 16-nm bulk FinFET technology nodes are simulated using the Monte-Carlo method in HSPICE. For 45-nm, the random effects are modeled through the variation in V_{th} caused by RDF and LER effects [46]. Similarly, for 16-nm, the variation is due to RDF, WKF, and ITs effects [31, 44]. The standard deviation σV_{th} values of 25.9mV for 45-nm process and 28.7mV for 16-nm process are adopted from [46] and [44], respectively. Finally, we restrict our discussion to show how these PV effects combine in redundant systems of logic datapaths to exhibit a higher mean delay than a simplex system. While numerous studies in the literature discuss recent innovations for SE-mitigation in memory devices, we focus herein on logic critical datapaths due to its increasing importance. The redundant system performance is determined by the worst-case delay out of any of the constituent modules. For instance, the delay of a TMR arrangement can be predicted to exceed that of any single module since the module instance with maximum delay determines its final delay. Generally, if the worst-case delay of module instance i of a redundant system is τ_i, then the overall delay of the TMR system τ_{TMR}, temporal system τ_{Temp}, and hybrid system τ_{SVDMR} are given by:

$$\tau_{_{TMR}} = \max_{1 \leq i \leq 3}(\tau_i) \; + \; \delta_{voter} \tag{4.1}$$

$$\tau_{_{Temp}} = \tau_i + \delta_{voter} + 2 * \delta_{SET} \tag{4.2}$$

$$\tau_{SVDMR} = \max_{1 \leq i \leq 2}(\tau_i) \; + \; 2 * \delta_{voter} + \delta_{SET} \tag{4.3}$$

Where δ_{voter} and δ_{SET} represent the delay of the voting logic and the delay of transient pulse width, respectively, which contribute directly to the critical delay. $2 * \delta_{SET}$ is required as a phase shift between CLK1 and CLK3, Figure 2.10, to ensure that the legitimate data is captured at the registers, whereas $2 * \delta voter$ is required in SV-DMR because one majority voter and one self-voter are located in the longest critical datapath [4].

4.4 Experimental Setup

4.4.1 Simulation Objectives, Tools and Workflow

Experiments are carried out to analyze the overheads for the above-mentioned soft error mitigation techniques in terms of energy consumption and the σV_{th} impact on the propagation delay. For this case study, an inverter chain composed of 26 Fanout-of-4 inverters with registered outputs was synthesized and simulated using 45-nm MOSFET planar and 16nm-gate High-κ/Metal Gate (HK-MG) bulk FinFETs structure based on PTM-based NanGate open source library [131], using an HSPICE simulation. Monte-Carlo simulations were carried out to implement the threshold variation in spatial, temporal, and hybrid (SV-DMR) redundancy schemes. These simulations vary the V_{th} of the transistor in the netlist based on a Gaussian distribution having a mean equal to the nominal model card for PTM and V_{th} as provided in [44, 46]. The σV_{th} can be adapted to accommodate local and global variations, or their combined effects as considered in this work.

First, operation at NTV increases the effect of threshold variation on the datapath delay of the soft error mitigation techniques redundancy-based. Therefore, using Eq. 4.1 through Eq. 4.3, we quantify the delay performance impact of threshold voltage variation for each mitigation technique.

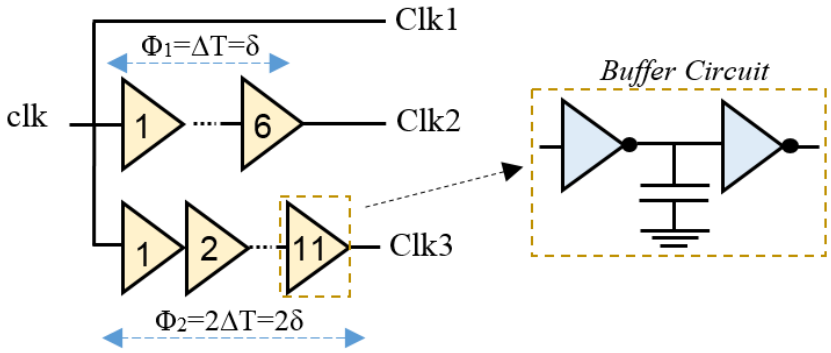

Figure 4.1: Local Clock Manager (LCM) to generate CLK2 and CLK3.

For example, Eq. 4.1 is used to determine the delay of a TMR system. The overall delay in this case is determined by the slowest critical datapath among the three modules and the delay of the voter circuit. On the other hand, there exists only a single datapath in the temporal redundancy approach while intricate details about the working of the voter circuit determine the overall delay as quantified in Eq. 4.2. The design of the temporal redundancy scheme is depicted in Figure 2.10, in chapter 2. It can be observed that the data is captured at three different time instances (T1, T2, and T3) by using three identical registers (flip-flops) triggered by three different clock signals (CLK1, CLK2, and CLK3). The relative latency between the clock signals is employed such that they are delayed by a phase shift (Φ_1 and Φ_2), where there is a phase delay Φ_1 between CLK1 and CLK2 and Φ_2 between CLK1 and CLK3. To further clarify, Figure 4.1 illustrates the internal working of Local Clock Manager (LCM) which is used to generate the required clock signals. Finally, the delay of ΔT can be selected depending on the SET pulse width coverage, thereby the total datapath delay is determined by summing the datapath and voter circuit's delay, while considering the slowest clock rate (CLK3) as delay of $2 * \delta_{SET}$ should be added.

In addition, the tradeoff analysis for the hybrid redundancy approach, Self-Voting DMR [4] (Figure 2.11), was also carried to for delay variation estimation. We calculate slowest datapath delay

among two modules, the original and the redundant module. However, it uses two voting circuits, one majority voter and one self-voter, therefore, $2 * \delta_{voter}$ is required for the longest delay datapath. In addition, it also incurs a phase delay within a period that depends on the SET pulse width to delay the clock rate of the third flip-flop, more details can be found in [4]. Thus, the total datapath delay is determined by summing the critical datapath and the delay of two voter circuits, while considering the slowest clock rate (CLK2) as quantified in Eq. 4.3.

The considered equations are leveraged in HSPICE circuit simulation tool to calculate the critical path delay for each scheme, and investigate the effect of threshold variation on each SER mitigation technique. The simulation framework is illustrated in 4.2. The Monte-Carlo simulations were conducted to utilize at least 1,000 experimental runs for every design implementation, therefore, the datapath delay of TMR system is obtained by calculating the mean value of the datapath delay of each module for 1,000 runs. Then the module instance with maximum datapath delay determines the delay of the TMR system as indicated in Eq. 4.1. Mean values are reported for each case. This scenario is repeated 1,000 times to establish mean value for every supply voltage value, as the supply voltage was altered from the nominal V_{DD} (1.1V) to 0.5V by a discernment of 0.5V for the 45-nm CMOS planar structure, whereas it is altered from the nominal V_{DD} (0.8V) to 0.5V for the 16-nm bulk FinFET structure. Furthermore, the energy consumption is computed by accumulating the energy requirement of the mean value operating at a frequency of $1/\tau_{TMR}$. The same setup is altered for other redundancy-based SE mitigation techniques presented in this work.

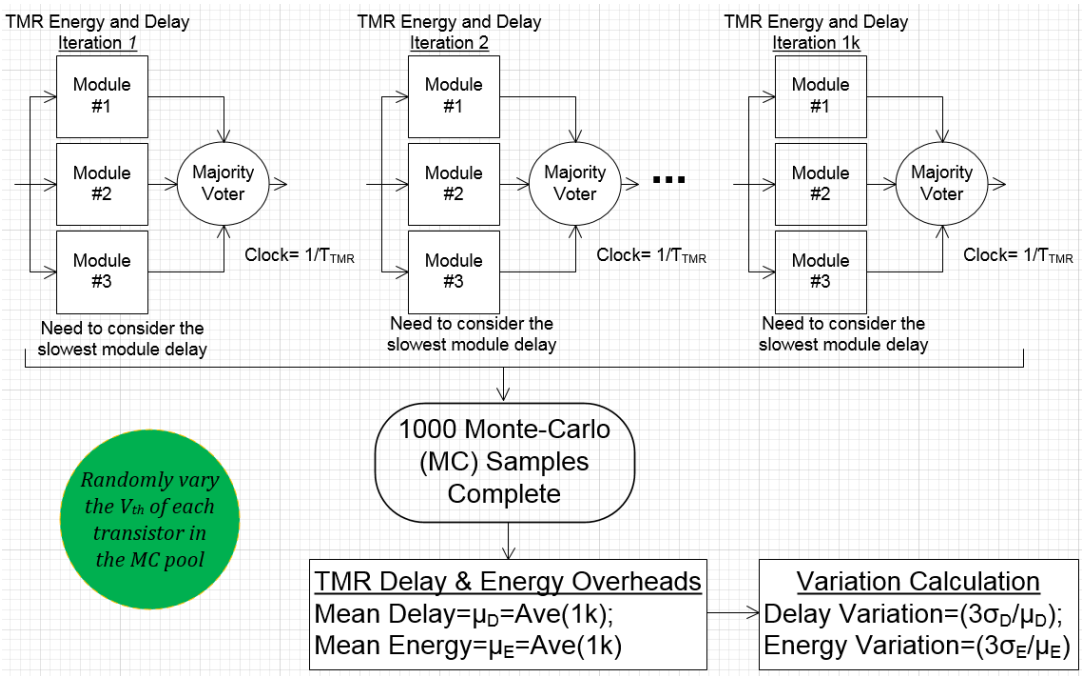

Figure 4.2: Simulation framework developed to estimate the delay and energy for TMR. At least 1,000 samples are synthesized for each redundant system to conduct the statistical analysis.

4.5 Experimental Results and Analysis

4.5.1 Quantifying Area Overhead for SE-Mitigation Techniques

It is known that the area overhead for the TMR is more than 200% including voting logic, whereas for temporal redundancy only the latches in the design grow in number while all the combinational logic elements remain unchanged. Thus, area overhead incurred can be expected to be roughly twice the area of the latches in the overall design. Indeed, a design which consists of a large sequential ratio and/or a large word-width occupies larger area as more flip-flops are required. On the other hand, area overhead of SV-DMR scheme presents a compromise between spatial and temporal redundancy approaches. For the fault coverage, the spatial, temporal, and hybrid redundancy

approaches achieve complete fault masking coverage for SEU in single module simultaneously. However, temporal and hybrid redundancy are capable of detecting and correcting all upsets occurring on registers, but they are unable to detect any transient pulse width exceeding 160 psec, as explained later on. Therefore, to maintain high coverage for all redundant systems in order to enable a fair comparison, a higher value of SET pulse width is considered herein than estimated in prior works [9].

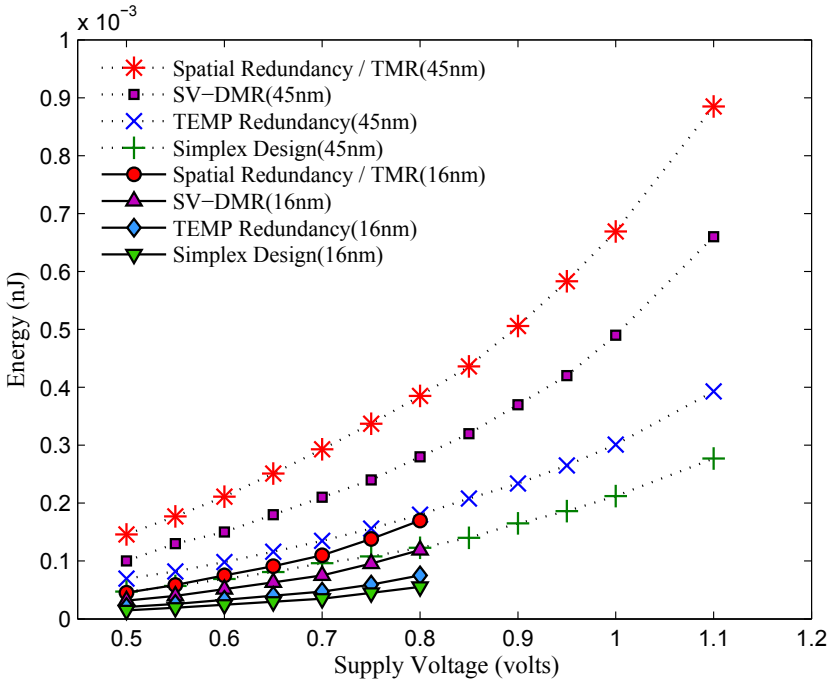

Figure 4.3: Mean energy consumption for redundant systems using 45-nm and 16-nm technologies.

4.5.2 Comparison of Delay and Energy Consumption

Results Figure 4.3 depict energy consumption for the simplex (non-redundant) and redundant arrangements. It is observed that temporal approach consumes a comparable amount of energy compared to the unprotected circuit. However, it incurs an average speed degradation of 28.19%

compared to TMR. In addition, as listed in Table 4.1, the experimental results depict an increment in the speed degradation for both temporal and hybrid redundancy approaches with scaled supply voltage. This is because buffering the clock signals causes more delay with scaling V_{DD}, buffering the clock signals will be discussed late in Section 4.5.5. Similarly, the speed degradation of both TMR and SV-DMR is seen to be produced from both considering the slowest critical delay path due the V_{th} variation and the delay of voter(s) circuit, besides considering the slowest clock rate for SV-DMR as quantified in Eq. 4.3. Figure 4.4 shows speed degradation analysis for redundant systems at NTV region using 45-nm technology process.

To further analyze the impact of increased variability on energy overhead of soft error resilient designs, experiments were conducted using the 16-nm PTM-MG HP model. As shown in Table 4.1, which illustrates the energy reduction and speed degradation for different supply voltages, the temporal redundancy realizes higher average energy saving as compared to TMR implementation, 54.14% and 56.21% for 45-nm and 16-nm, respectively. This is partly due to the fact that the multi-gate FinFET devices have less leakage than the planar MOSFET devices [30, 31]. In term of speed performance, the speed degradation, Figure 4.5, of both temporal and hybrid redundancy is reduced while utilizing multi-gate bulk FinFET devices because the latter achieves more robust gate controllability and thus improves the system performance. Thus, both temporal or hybrid redundancy approaches are able to exploit the benefits of scaled technology devices. The overall comparison of energy and delay for all redundant system is presented in Table 4.2. On average, temporal scheme consumes 34.51% more energy than the simplex circuit, whereas the energy overhead is 207.19% and 123.25% for TMR and SV-DMR, respectively. However, temporal redundancy degrades system performance by 37.79% within a $\delta_{SET} = 150$ psec for transient pulse width using 16-nm technology.

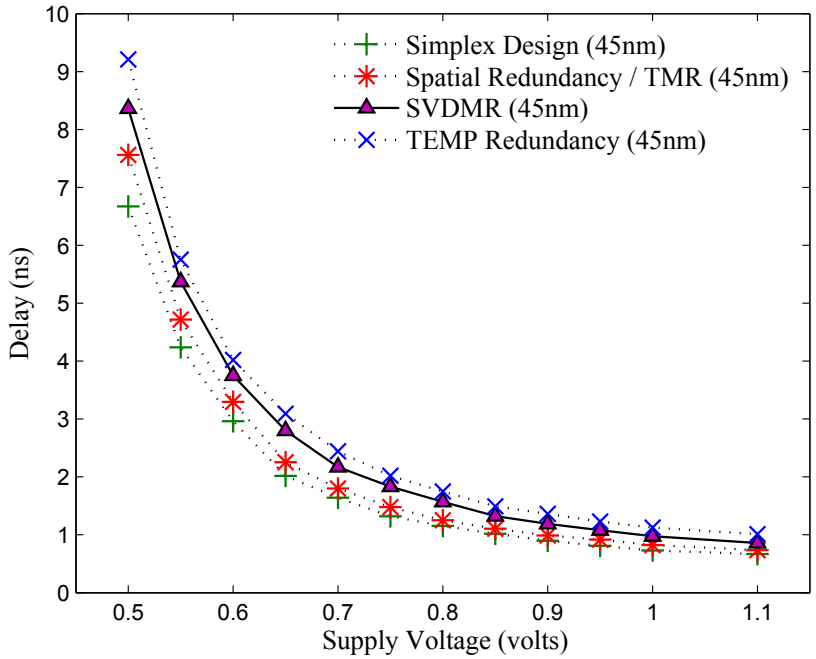

Figure 4.4: Speed degradation analysis at NTV for the 45-nm technology.

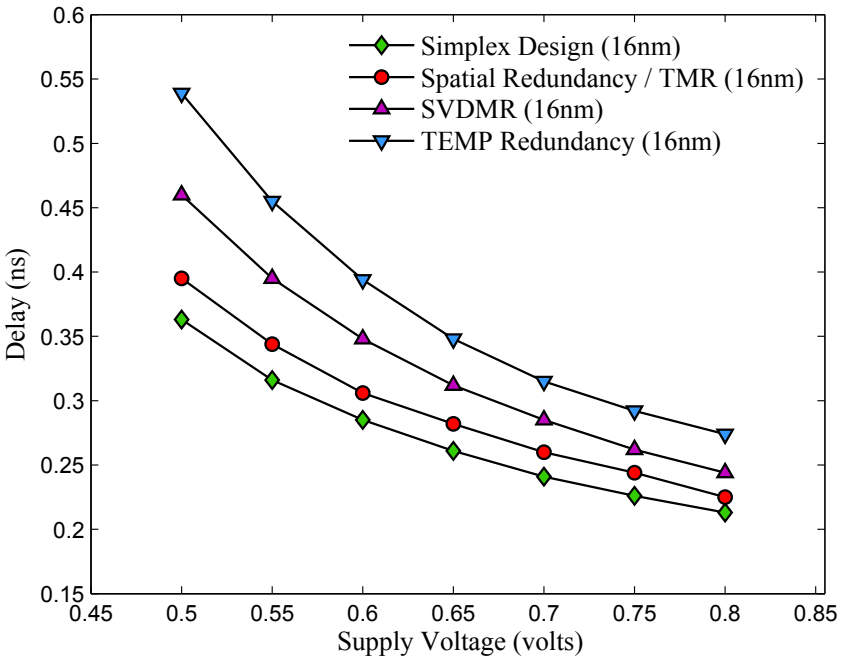

Figure 4.5: Speed degradation analysis at NTV for the 16-nm technology.

Table 4.1: Mean energy reduction and speed degradation of temporal redundancy and SV-DMR approach versus TMR.

V_{DD}	45nm				16nm			
	Temporal		SV-DMR		Temporal		SV-DMR	
	Energy Reduct.	Speed Degrad.	Energy Reduct.	Speed Degrad.	Energy Reduct.	Speed Degrad.	Energy Reduct.	Speed Degrad.
$1.1V$	55.6%	21.22%	24.22%	12.92%	-	-	-	-
$0.8V$	54.41%	26.73%	23.67%	15.78%	55.15%	21.78%	27.31%	8.79%
$0.65V$	53.9%	31.9%	23.25%	18.49%	55.86%	26.24%	26.85%	12.87%
$0.55V$	53.6%	33.16%	22.18%	21.53%	55.17%	30.27%	26.53%	15.34%
$0.5V$	52.45%	35.76%	21.54%	23.17%	54.57%	33.45%	25.84%	17.74%
Ave.	54.14%	28.19%	22.31%	16.6%	56.21%	27.67%	26.93%	15.86%

Table 4.2: Mean energy consumption and speed degradation of spatial, temporal, and hybrid redundancy approaches.

Design Implementation	45nm		16nm		vs. Simplex 16nm	
	Energy (pJ)	Speed (ns)	Energy (pJ)	Speed (ns)	Energy	Delay
TMR	0.391	2.26	0.0932	0.294	307.19%	7.92%
Temporal	0.191	2.93	0.0464	0.374	134.51%	37.79%
SV-DMR	0.296	2.595	0.0701	0.334	223.25%	24.94%

4.5.3 Impact of Process Variation in NTV Region

Table 4.3 depicts the speed degradation in the NTV region of redundant systems normalized to the simplex design implementation. It can be seen that the speed degradation exacerbates with lowering V_{DD} for redundant implementations. This implies that the variation in the output delay was found to be increased under scaling down V_{DD} from the nominal level to 0.5V for both the 16-nm bulk FinFET and 45-nm planar MOSFET. The critical path and SET pulse width for spatial and temporal redundancy respectively incur more speed degradation with down scaling of the supply voltage. As illustrated in Figure 4.6, the delay variation of temporal redundancy is lower than the variation of both TMR and SV-DMR, even the variation of 16-nm is beneath that of 45-nm

technology node for spatial and hybrid redundancy approach. This implies that delay variation increases for either higher redundancy or increased sophistication of the fault resolution circuit. However, the latter impacts the delay variation higher as can be seen in the curves of SV-DMR approach in Figure 4.6 for both 45-nm and 16-nm technology node. Thus, temporal redundancy can be utilized to alleviate the effect of delay variation at NTV. In addition, the 16-nm node-based temporal system improves the normalized speed degradation by 12.4% compared to the implementation of a temporal system based on 45-nm planar, and it also enhances the average normalized energy saving by 5.74% compared to the 45-nm planar structure. Consequently, soft error resilient low power designs can be protected more adequately by employing the temporal approach at lower V_{DD}.

Table 4.3: Normalized speed degradation of spatial, temporal, and hybrid redundancy (SV-DMR) w.r.t. simplex design.

$\mathbf{V_{DD}}$	**45nm**			**16nm**		
	Spatial Redund.	**Temporal Redund.**	**Hybrid (SV-DMR)**	**Spatial Redund.**	**Temporal Redund.**	**Hybrid (SV-DMR)**
$0.8V$	1.105x	1.33x	1.25x	1.05x	1.29x	1.11x
$0.75V$	1.107x	1.34x	1.27x	1.05x	1.3x	1.13x
$0.7V$	1.109x	1.36x	1.28x	1.06x	1.32x	1.13x
$0.65V$	1.11x	1.37x	1.3x	1.07x	1.34x	1.14x
$0.6V$	1.11x	1.39x	1.31x	1.07x	1.35x	1.15x
$0.55V$	1.12x	1.42x	1.33x	1.09x	1.36x	1.17x
$0.5V$	1.13x	1.44x	1.35x	1.09x	1.38x	1.17x

4.5.4 Impact of Supply Voltage in NTV Region

NTV operation reduces supply voltage to 100-200mV above the threshold voltage of the transistors, allowing for significant improvement in energy-efficiency with a reasonable performance impact [122]. The energy savings can be utilized for either reduced power consumption or to increase

resilience via redundant implementation. Thus, the near-threshold region allows for consideration of interesting tradeoffs. For example, a design with spatial or temporal redundancy can be utilized as a means to increase reliability within the same energy budget as a simplex (unprotected) system operating at nominal voltage. This is valid provided that the increase in delay, and thus corresponding drop in performance and area costs are acceptable. Note that pursuit of increased reliability is predicated upon the assumption that the source of variability in the near-threshold region is due to variation in V_{th} for which this work is restricted.

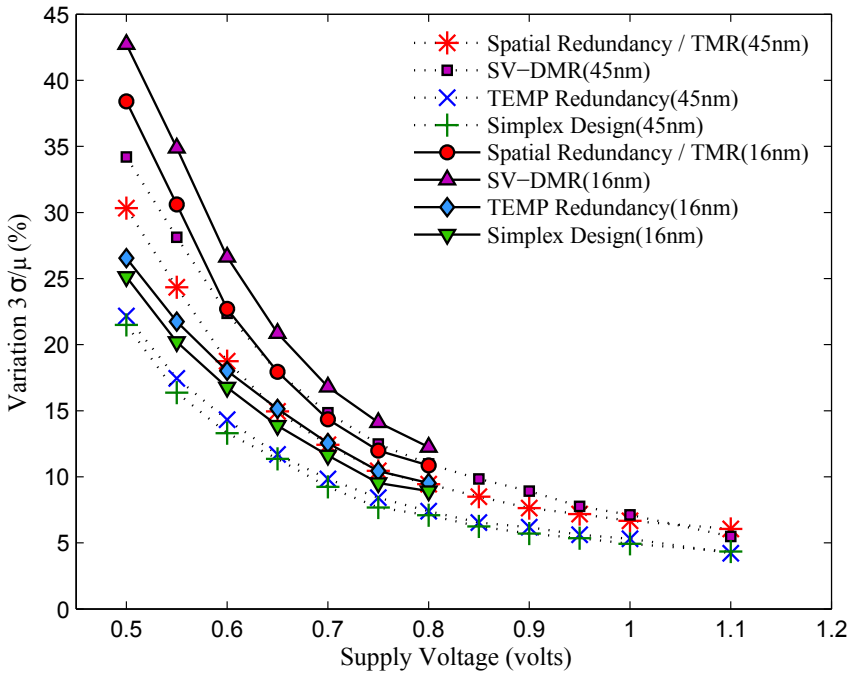

Figure 4.6: Delay variation for redundant systems using 45-nm and 16-nm technologies.

Based on these assumptions, the tradeoffs in Figure 4.7 and 4.8 are highlighted. For instance, it shows the feasibility of the temporal approach at around 0.97V on average given an identical energy budget of a simplex system operating at nominal voltage of 1.1V (for 45-nm technology). TMR and hybrid redundancy can be employed to protect the design at lower supply voltage, ~ 0.67V

and $\sim$ 0.8V, respectively. On the other hand, for 16-nm technology process, results emphasize that utilizing temporal redundancy approach at NTV provides better energy reduction with scaling, as shown in Figure 4.8. The temporal scheme can be employed to mitigate soft errors with the nominal energy budget at a supply voltage between 0.7V to 0.75V. Meanwhile, spatial and hybrid redundancy should operate between 0.5V to 0.55V, or 0.6V to 0.65V, respectively, to maintain the energy consumption at the same level as the simplex design implementation. In this case, using either TMR or SV-DMR in NTV region will degrade the performance more than utilizing temporal redundancy as shown in Table 4.4, where it depicts performance degradation with constant energy budget for redundant systems compared to simplex design implementation. Thus, while maintaining energy-efficiency temporal redundancy scheme provides better performance in terms of energy saving, speed degradation, and variation in output delay with technology scaling at NTV. Therefore, for applications which seek to protect a design against soft errors with constant energy dissipation budget as compared to nominal operation, it is advantageous to utilize temporal redundancy approach to achieve that at lower V_{DD} since scaling down the supply voltage is the most effective method to reduce energy consumption.

Table 4.4: Effect of reducing supply voltage under iso-energy constraints.

Design Implementation	**45nm**			**16nm**		
	V_{DD} (volts)	**Norm. Energy**	**Norm. Delay**	**V_{DD}** (volts)	**Norm. Energy**	**Norm. Delay**
Simplex	1.1	1x	1x	0.8	1x	1x
TMR	$\sim$ 0.67	1.059x	2.7x	0.5 $\sim$ 0.55	1.057x	1.615x
Temporal	$\sim$ 0.97	1.087x	1.68x	0.7 $\sim$ 0.75	1.072x	1.37x
SV-DMR	$\sim$ 0.8	1.037x	2.35x	0.6 $\sim$ 0.65	1.014x	1.48x

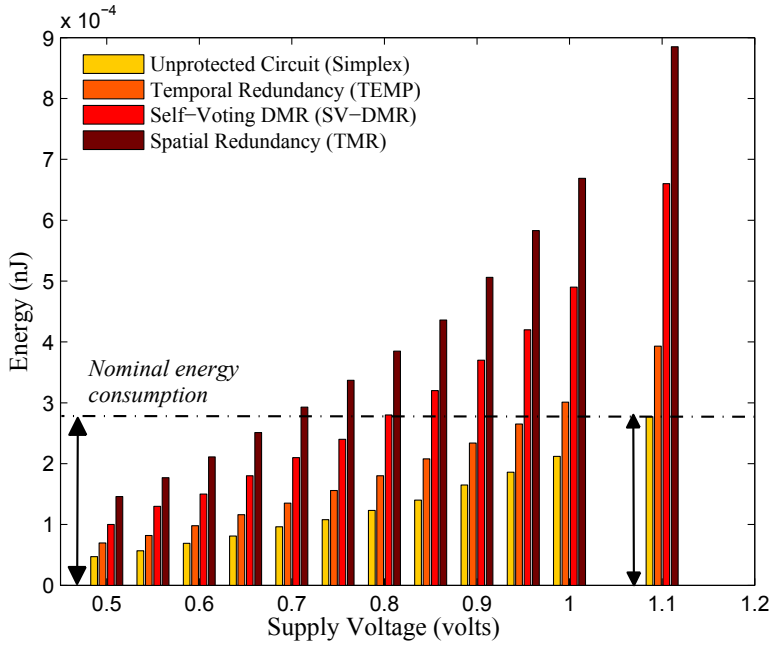

Figure 4.7: Maintaining energy consumption at NTV using 45-nm technology process.

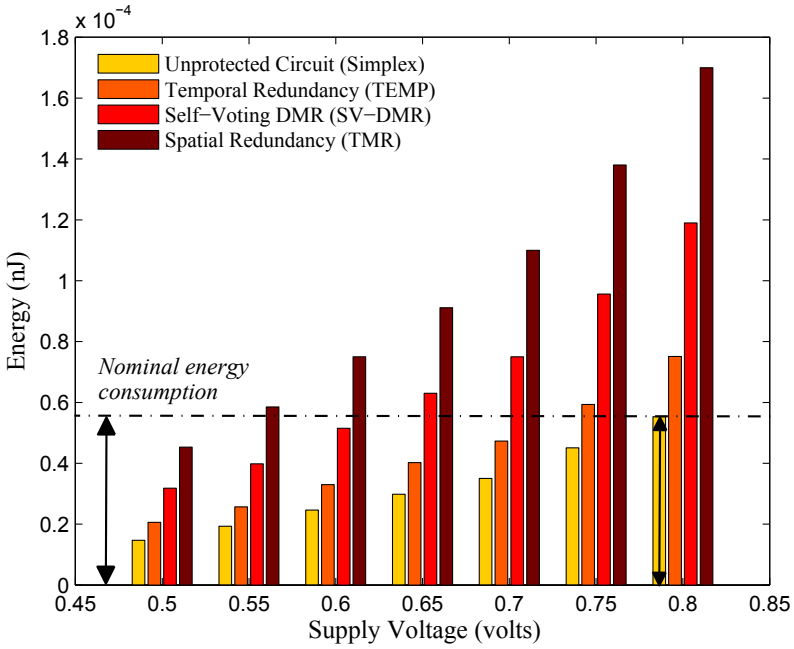

Figure 4.8: Maintaining energy consumption at NTV using 16-nm technology process.

4.5.5 Constraints of Temporal Redundancy Approach

Herein, experiments are conducted to explore the tradeoffs of performance vs SET pulse width for the temporal approach. To generate the clock signals for temporal redundancy approach, shown in Figure 2.10, a Local Clock Manager (LCM) is employed, presented in [85], offering advantages in terms of complexity and energy consumption. CLK1 is the same as CLK (main clock), while CLK2 and CLK3 are produced by utilizing clock buffers. The number of buffers depends on the phase shift between the clocks. The latter depends on SET pulse width. The transient pulse width is estimated in [9, 132] with nominal voltage operation and 65-nm technology node between 25 psec to 125 psec at the sea level. In this work, a relatively higher pulse width value of 160 psec and 150 psec was chosen to accommodate the worst case and incorporate the effect of using reduced technology node size of 45-nm and 16-nm, respectively.

As illustrated in Figure 4.9, the considered delay, δ_{SET}, increases as the supply voltage is lowered. Thus, operation at NTV increases the delay for the clock buffers, where the number of buffers were determined based on nominal value of V_{DD}, and thus, this makes the temporal redundancy scheme more robust to reject larger SET pulse width than operation at nominal voltage. Overall, the temporal redundancy approach is able to reject SET pulse widths ranging between 380 psec to 850 psec and 170 psec to 230 psec when the supply voltage is scaled between 0.7V to 0.55V, for 45-nm and 16-nm technology process, respectively. This demonstrates the benefit of temporal approach for these parameters. However, this is at the expense of performance degradation as stated in Section 4.5.2.

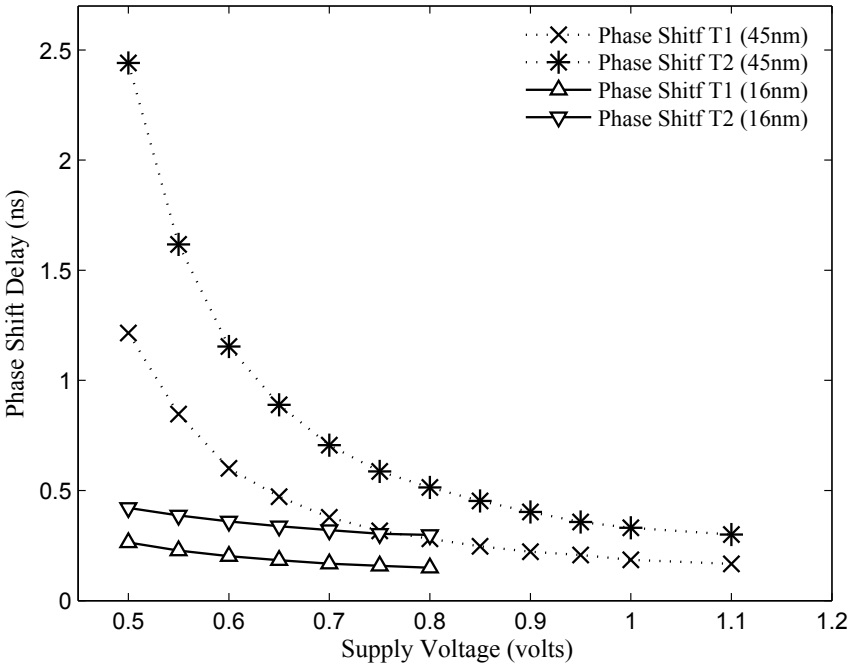

Figure 4.9: SETs pulse width rejection at NTV for 45-nm planar MOSFET and tri-gate 16-nm bulk FinFET.

4.6 Energy Efficency versus Fault Resiliency

To realize tradeoffs between the fault-masking coverage achieved versus its cost in terms of power and delay, an appropriate metric is proposed. The metric is then used to optimize the design of several low-power circuits. We evaluate the redundant arrangements by calculating a new metric called *Fault Coverage Energy Ratio (FCER)*, which divides the soft or transient fault masking coverage over the worst case energy consumption. The aim is to maximize the resilience of fault coverage and minimize the consumed energy. Thus, the higher the value of FCER metric, the

better the protection arrangement. FCER is expressed as:

$$FCER = \frac{fault\ coverage}{MAX\ (delay \times power\ consumption)} \tag{4.4}$$

The worst case energy consumption considers the power consumed by the logic circuit as well as the protection circuit. To consider worst case conditions, the circuit needs to be evaluated through the application of worst case input vectors. FCER is used to evaluate the overheads of providing soft error resilience using multiple mitigation techniques. We identify the following applications of FCER metric: 1) FCER serves to compare multiple mitigation techniques; 2) Use of FCER as a parameter for synthesis of resilient circuits. In an iterative design process, the metric can be used as an optimization parameter to guide selection of the protection approach and meet performance and energy goals.

4.6.1 FCER Evaluation of Redundancy-based Approaches

The summary of the FCER analysis for the redundancy-based soft error mitigation techniques is listed in Table 4.5. Note that, only the fault masking coverage for TSVL approach is listed in Table 4.5, while the masking coverage for other redundancy-based techniques is assumed to be 100%. Additionally, Figure 4.10 shows FCER under scaling down V_{DD} from the nominal level (1.1V) to 0.5V for 45nm technology. As observed in Table 4.5, at the nominal supply voltage, the hybrid redundancy scheme (TSVL) achieves the preferred tradeoffs in terms of FCER among the redundancy-based soft error mitigation approaches for most of the selected benchmark circuits, irrespective of whether the combinational or the sequential logic ratio is higher. While the temporal redundancy realizes high or acceptable FCER when the combinational logic ratio is high, TMR achieves desirable/acceptable FCER either when the sequential logic ratio is high or when

the benchmark circuit occupies a small area. On the other hand, in the NTV region, the temporal redundancy approach realizes an increased benefit in terms of FCER, meanwhile TSVL still achieves competitive FCER compared to the temporal approach at NTV operations. Consequently, the TSVL approach can be utilized to harden logic paths against radiation-induced soft errors that are constrained with an iso-energy consumption budget with a minimal impact on the masking coverage roughly by 1.5% of transient errors. Likewise, the temporal approach offers promising attributes at NTV operation while encountering an acceptable/reasonable performance delay variation and degradation.

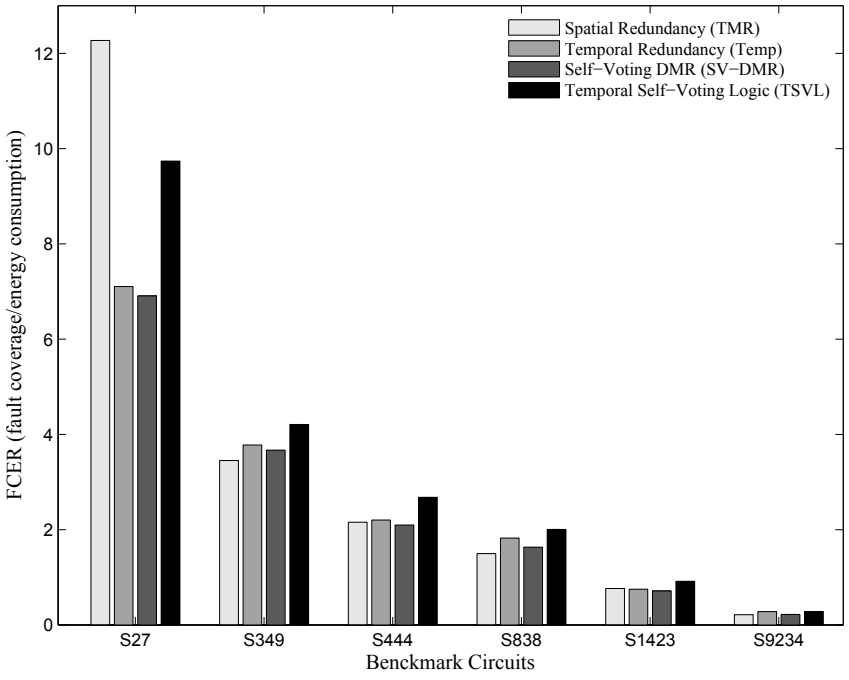

Figure 4.10: Mean fault coverage energy ratio analysis for benchmark circuits realizing different redundancy-based approaches. Each point is obtained by average of FCERs obtained by scaling V_{DD} from nominal level (1.1V) to the NTV region (0.5V) with decrements of 0.1V.

Table 4.5: Fault coverage energy ratio analysis of redundancy-based soft-error mitigation approaches.{FC=Fault Coverage}

Benchmark Circuit	**♯ of Cells**	**FC of TSVL** (%)	**Redundancy Approach**	**FCER** (V_{DD}=1.1V)	**FCER** (V_{DD}=0.75V)	**FCER** (V_{DD}=0.55V)
$S27$	13	97.4	TMR	12.82	12.25	10.28
			Temporal [2]	7.11	6.88	7.47
			SV-DMR [4]	7.034	6.97	5.45
			TSVL [67]	9.90	9.99	7.99
$S349$	127	97.075	TMR	3.60	3.4467	2.8929
			Temporal	3.77	3.6617	3.9745
			SV-DMR	3.73	3.7022	2.8993
			TSVL	4.26	4.3210	3.4546
$S444$	163	97.34	TMR	2.25	2.1523	1.8065
			Temporal	2.19	2.1330	2.3153
			SV-DMR	2.13	2.1165	1.6574
			TSVL	2.71	2.7488	2.1976
$S838$	418	97.11	TMR	1.55	1.4937	1.2537
			Temporal	1.82	1.7662	1.9172
			SV-DMR	1.66	1.6468	1.2896
			TSVL	2.027	2.0595	1.6466
$S1423$	683	96.84	TMR	0.798	0.7630	0.6404
			Temporal	0.748	0.7247	0.7866
			SV-DMR	0.724	0.7179	0.5622
			TSVL	0.924	0.9392	0.7509
$S9234$	1369	96.24	TMR	0.223	0.2140	0.1797
			Temporal	0.278	0.2690	0.2920
			SV-DMR	0.224	0.2219	0.1738
			TSVL	0.282	0.2888	0.2309
Simulation Results of Complete Systems Utilization						
32-bit Pipeline Processor	22360	94.34	TMR	0.0686	0.0812	0.0524
			Temporal	0.0759	0.1003	0.0654
			SV-DMR	0.0587	0.0843	0.0546
			TSVL	0.0760	0.0.975	0.0843
128-bit AES Encryption	12828	95.61	TMR	0.0159	0.0189	0.0122
			Temporal	0.0311	0.0410	0.0268
			SV-DMR	0.0185	0.0267	0.0173
			TSVL	0.0267	0.0.342	0.0296
I^2C Memory Bus Controller	1095	93.15	TMR	0.0427	0.0507	0.0327
			Temporal	0.0771	0.1018	0.0664
			SV-DMR	0.0539	0.0775	0.0502
			TSVL	0.0653	0.0.838	0.0725

4.6.2 Design and Analysis using FCER

To further investigate the energy-efficiency and resilience of fault masking coverage for highlighted approaches, we synthesized and optimized via FCER maximization the following applications in addition to the benchmark circuits: 1) a 5-stage pipeline processor, 2) a cryptographic processing core, and 3) a memory controller. Using the results, realistic estimations are obtained of performance penalties versus robustness for mission-critical systems. These circuits were synthesized and simulated using Synopsys Design Compiler with the 45nm technology node of NanGate open source library. We show how energy savings obtained from NTV operation can be utilized to fortify both the computational logic and the pipeline registers against soft errors via the aforementioned mitigation approaches. Then, we investigate the effects of soft errors on the protected computational logic while evaluating area overheads. For example, a TMR arrangement is utilized with majority voting which has the ability to mask corrupted outputs. On the other hand, a low overhead approach as compared to TMR, the Self-Voting DMR [4] approach is also evaluated. Alternatively, the temporal and hybrid (TSVL) redundancy approach replaces the pipeline registers so that soft fault detection and correction storing circuitry is deployed only in the pipeline registers while utilizing a simplex instance of the datapath.

The simulation results of a 5-stage pipeline 32-bit MIPS processor show that the hybrid redundancy (TSVL) scheme incurs only an increased energy of 5.7% at nominal supply voltage compared to the temporal redundancy approach, which incurs the lowest power consumption and area overhead. This is due to a major contribution of TSVL hardening scheme as it utilizes a single computational logic for error detection and duplicated registers, instead of triplicated registers in the temporal redundancy arrangement. Therefore, performance penalties of TSVL are reduced compared to the conventional TMR or SV-DMR pipeline stages. However, for the Advanced Encryption Standard (AES) encryption circuitry, temporal redundancy realizes the best energy results among all uti-

lized techniques. This is because the AES encryption module has large fraction of computational logic that remains unprotected in the temporal redundancy while only the state holding logic, i.e., registers/flip-flops, are triplicated. Therefore, a minimum area overhead is required to immune the registers. Figure 4.11 and 4.12 demonstrate that the temporal redundancy delivers an optimized Energy Delay Product (EDP) at 0.8V, whereas TSVL needs to operate at 0.75V to realize the optimal EDP. Even though temporal redundancy provides reduced energy savings, it has higher propagation delay at the NTV region due to the buffer circuits which are used to delay the main CLK signal and generate the phase shifted signals to capture the input data at different time instances. Meanwhile, considering low supply voltage (V_{DD}=0.75 or 0.55V) for energy-efficiency, the TSVL approach delivers intriguing energy savings. This competes well with the temporal redundancy approach by achieving roughly 2.96% and 26.76% improvement, respectively. Therefore, it can be utilized to fortify low-power circuits and systems against transient and upset faults. On the other hand, both the conventional TMR and SV-DMR redundancy approaches incur high energy and EDP compared to either the temporal or TSVL approaches at NTV. Consequently, variation-aware fault-tolerant designs can be hardened via temporal redundancy. This is because it exhibits strong capability to incur less delay variation [29] in the critical output datapath and preserves the best tradeoff between energy saving and fault masking coverage for operations at the nominal supply voltage, as listed in Table 4.5 and discussed later on.

4.6.3 Evaluation of Fault Coverage

In this work, we utilize an in-house fault-injection module developed at the RTL-level to simulate and evaluate the transient fault masking capability of circuits implemented in Verilog HDL. Large-sized circuits such as a 32-bit pipeline processor, a 128-bit AES encryption core, and an I^2C controller have been evaluated. This aids us to accelerate the determination of FCER for practical systems.

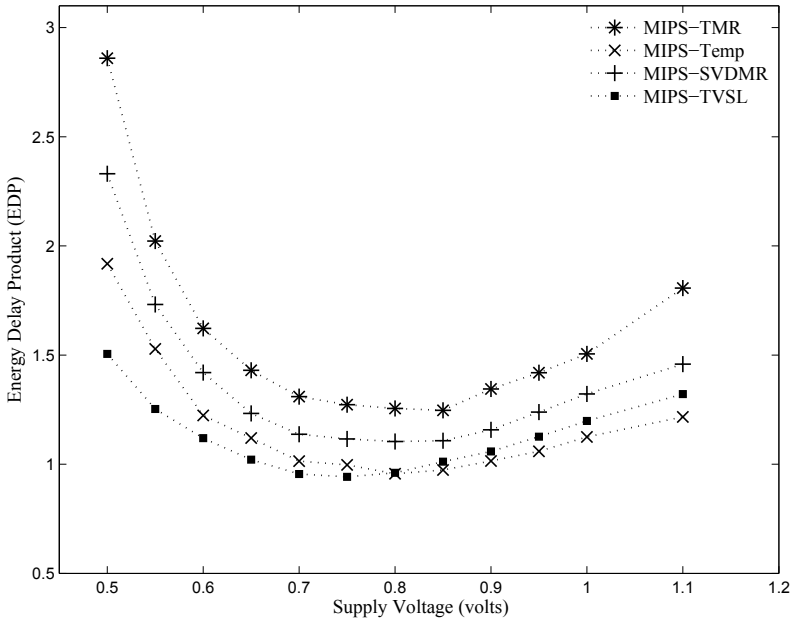

Figure 4.11: Energy Delay Product (EDP) analysis with varying supply voltage for multiple redundant arrangements using a 32-bit 5-stage pipelined MIPS processor.

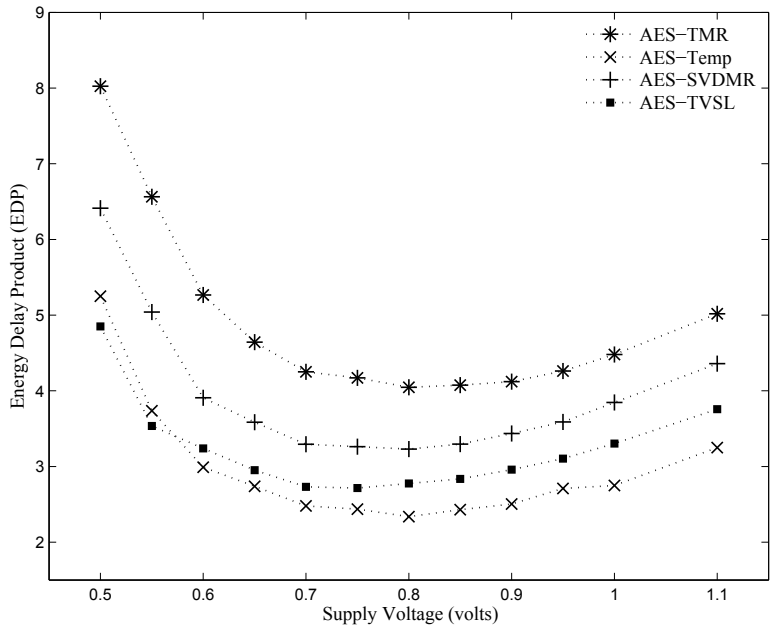

Figure 4.12: Energy Delay Product (EDP) analysis with varying supply voltage for multiple redundant arrangements using a 128-bit Advanced Encryption Standard (AES) algorithm.

For experimental results herein, 1,000 faults of each SET and SEU were injected into the gate-level Verilog code describing the benchmark circuitry configured with the redundancy-based mitigation techniques. SETs were injected into a node impacting a logic gate, while SEUs were injected into the input of a register/flip-flop with the rising edge of the main CLK signal. All errors were injected at randomly selected locations. The evaluation results showed that the spatial (TMR) and hybrid (SV-DMR) redundancy can prevent the injected transient fault from propagating through the computational logic of the pipeline stages in all cases. On the other hand, for the hybrid redundancy (TSVL), 5.66% of the injected errors were occasionally observed to corrupt a pipeline register causing transient error in the register's output.

Further investigation into the fault coverage results of the TSVL approach listed in Table 4.5 reveals interesting trends. Specifically, TSVL approach achieves high fault resilience for systems that include higher proportion of combinatorial logic than sequential logic. For instance, the majority of the logic (roughly 85%) of the AES encryption circuitry is combinational logic for computation while only about 15% is used for registers. Therefore, in this case, the possibility of having a SEU at the intermediate registered output hardened with the hybrid redundancy is reduced. On the other hand, for the I^2C controller benchmark circuit, the TSVL approach provides inadequate robustness (93.15%) since the I^2C controller circuitry includes high sequential logic portion for holding its state, and thus there is a higher probability for an intermediate register to capture a SET and cause an upset. Even with these unpreventable transient faults, the TSVL approach still achieves a high soft fault masking with roughly 7% margin of error.

4.6.4 Summary of Fault-tolerant Design using FCER

The peak tradeoffs between energy efficiency and resilience of fault masking for the hybrid (TSVL) compared to temporal redundancy approach can be obtained at NTV (V_{DD}=0.75V), as an energy

saving of 2.96% is achieved. Furthermore, the hybrid redundancy approach realizes a higher FCER than the temporal redundancy as shown in Table 4.5 and Figure 4.13. However, our results demonstrate that this is worthwhile for an application/system with the following traits: balanced sequential and combinational resources or an application constrained with iso-energy objective that allows sacrificing fault resilience while meeting the constrained budget of energy consumption. For instance, image processing applications are a promising field to employ the hybrid redundancy at NTV region. This demonstrates the benefit of utilizing either the TSVL or temporal redundancy approach at NTV to protect low power designs against transient faults.

In summary, the energy usage is minimal while using the hybrid redundancy (TSVL) for protection, with the expense of delivering a reduced performance due to the speed degradation. Therefore, to improve reliability of low power computing systems in NTV region, the temporal redundancy approach can be selected to harden designs that consist of higher sequential logic resources, whereas the hybrid redundancy can be chosen to protect a design configured with a balanced usage between the computational and sequential logic resources. Tradeoffs between energy-efficiency and resilience of fault masking coverage can be obtained by utilizing the hybrid redundancy approach to protect a low-power system that requires maintaining high performance and tolerates few unmasked soft errors. Finally, the temporal redundancy approach can be selected to effectively immune mission-critical systems to radiation-induced transient faults while delivering energy-efficiency and preserving high soft fault masking with an acceptable performance degradation. In this case, FCER metric is validated as an optimization parameter to guide circuit synthesis algorithms to meet performance goals after the selection of a mitigation technique. It is noteworthy that the electrical and logical masking mechanisms that inherently mask transient faults are not considered during the fault injection process, thus, fault masking coverage can be expected to be higher in the field than the values estimated.

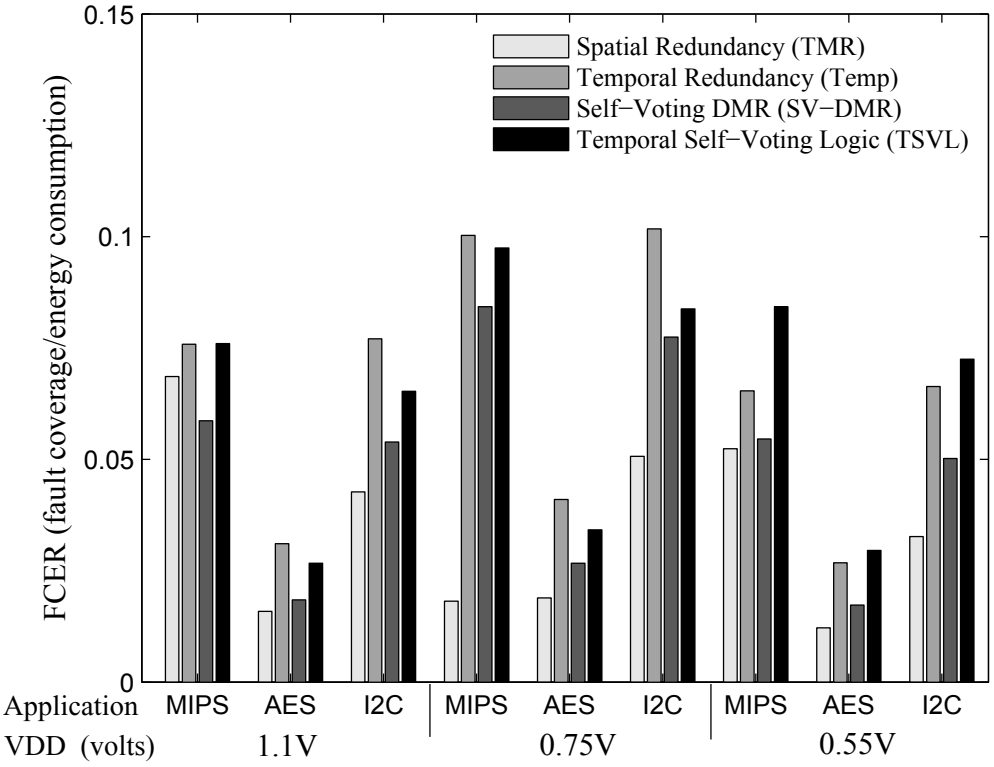

Figure 4.13: Fault coverage energy ratio analysis of multiple application circuits protected using multiple redundancy-based approaches, V_{DD} is scaled from nominal level (1.1V) to the NTV region (0.55V).

4.7 Relationship of Area and FCER

Herein, we explore the relationship between the SER and area overhead of redundant systems by using a set of ISCAS89 benchmark circuits with various number of cells. It turns out that the ratio for all circuits is the same except that the power is positively correlated with area and the probability of soft error is positively correlated with area. Thus, there is an overall impact of $FCER = O(Area^2)$ relationship as observed in Figure 4.14. Meanwhile, since the TSVL approach offers a reduced SER while incurring modest performance penalties in terms of occupied area, power consumption, and speed degradation, it is considered an intriguing approach that addresses or alleviates the drawbacks of spatial and temporal redundancy approaches.

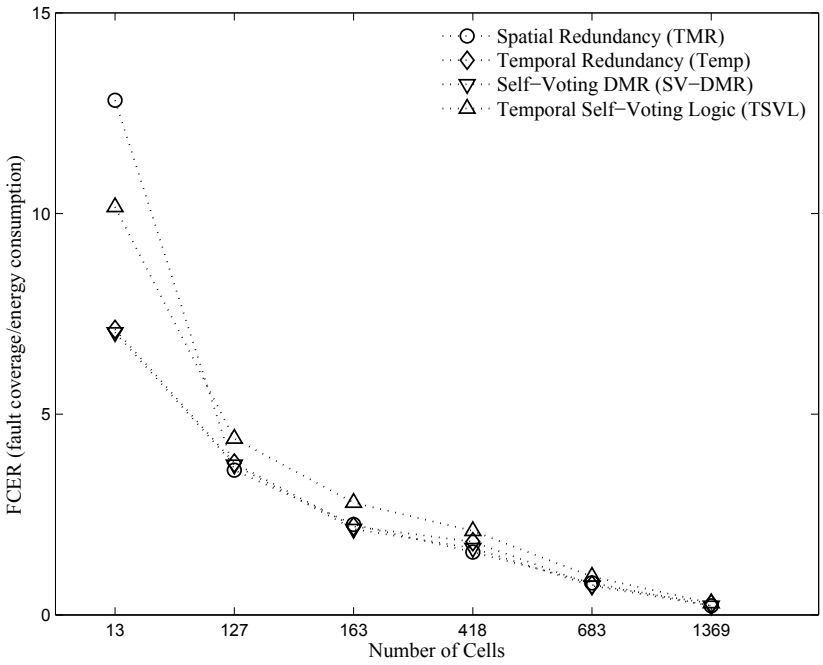

Figure 4.14: Relationship between area (number of cells) and FCER for each benchmark circuit realized using multiple redundancy approaches, at nominal supply voltage ($V_{DD} = 1.1V$).

4.8 Design Diversity and Spatial Reduandancy

Here, the use of Design Diversity, which is a circuit synthesis technique, is investigated to mitigate soft errors. Design diversity with spatial redundancy as a use-case is employed to test the efficacy of fault-tolerant circuit synthesis using FCER. For instance, a TMR arrangement is constructed with three different structures while a voter circuit determines the majority condition in order to determine the final masked output. To quantify the advantages/disadvantages of diverse TMR arrangements, the FCER metric defined earlier in Section 4.6 is utilized. Table 4.6 lists the summary for some design arrangements of TMR systems. Results indicate that the inverter-based chain achieves the highest FCER, but it incurs the highest delay variation to impact the performance as compared to diverse implementations. Additionally, these results are based on a single error in-

jection scenario considered for each diversity enabled-TMR inverter chain, since the conventional TMR approach is only able to tolerate a single error at a time. Thus, the diversity enabled-TMR arrangements might be used to alleviate/address the issue of MBU, aside from its benefit to tolerate the CMFs. Meanwhile, the delay variation results in time-dependent bit error, when the delay variation exceeds the clock period. Therefore, considering the transient error (SET/SEU) and timing error when calculating the fault masking coverage will provide a more accurate evaluation for the FCER metric. It provides a means to further highlight the pros and cons of design diversity when used with spatial redundancy.

Table 4.6: Fault coverage energy ratio analysis of diverse-TMR arrangements.

Diverse-TMR (Inverter Chain)	**Delay (ns)**	**Delay Variation** 3σ **(%)**	**FCER**
INV,INV,INV	0.61	2.46	3.94
INV,INV,NAND2	0.62	2.01	3.84
INV,NAND2,NAND2	0.745	1.56	3.158
NAND2,NAND2,NAND2	0.75	1.44	3.1

4.9 Summary

Mitigating soft errors at NTV can provide a range of alternatives across metrics of area, speed, and power. Thus, it would be advantageous to consider NTV operation within contemporary constraints of the design, such as minimum energy within a soft error mitigated design, maximum speed given an energy budget, or a tradeoff between these issues. In terms of delay variation, temporal redundancy incurs less variation (22%) at NTV ($V_{DD} = 550$mV) under technology node scaling as compared to both TMR and SV-DMR (31.6% and 35.2%, respectively), even the variation of 16-nm is beneath that of 45-nm technology node for both. Thus, the drawback of TMR and SV-DMR is the need to accommodate the slowest module delay. On the other hand, temporal

redundancy provides higher energy saving, but it requires consideration of the SET pulse duration. Thus, for soft error resilient low power designs, designers can select temporal redundancy at lower supply voltage thereby allowing for significant improvement in energy-efficiency at NTV under technology generation, while facing an acceptable delay variation and a reasonable speed degradation.

CHAPTER 5: HIGH-PERFORMANCE DOUBLE NODE UPSET-TOLERANT NON-VOLATILE FLIP-FLOP DESIGN

This chapter discusses implications for future emerging spin-based latching circuits. Recently, emerging spin-based devices are introduced as an intriguing candidate to alleviate leakage currents and continue the scalability of CMOS technology. However, their immunity to radiation-induced transient faults needs to be adequately addressed. In this work, a radiation-immune hybrid Spin Transfer Torque Magnetic Tunnel Junction (STT-MTJ)/CMOS flip-flop is designed and evaluated for nonvolatile applications. The proposed nonvolatile flip-flop circuit achieves attractive features, such as low standby power dissipation (21% less than CMOS-based design), high computing performance, and superior soft-error resilience (concurrently can tolerate DNU) to potentially become as a mainstream solution for the aerospace and avionic nanoelectronics.[1]

5.1 Hybrid CMOS/Spintronics: Alleviating CMOS Challenges

Currently, CMOS logic designs are moving toward their physical limits, and therefore, their continued scalability to sustain Moore's law becomes challenging in near future [134]. In addition, static power in CMOS logic and memory circuits of technologies beyond 45nm has increased due to increasing leakage currents [135]. To address this issue, a hybrid STT-MTJ/CMOS circuit design is developed as a potential alternative and an intriguing candidate to address scalability challenges. Moreover, the property of Non-Volatility (NV) facilitates the capability of instantly turning on devices that are normally OFF. Thus, the energy consumption demand can be reduced. Due its efficient recall operation and compatibility with CMOS processing, the STT-MTJ NV flip-flop has been targeted for nonvolatile processors. In particular, it is an intriguing concept to address power

[1] © 2018 IEEE. Reprinted, with permission, from [133].

failures from energy harvesting [136, 137].

An MTJ cell is inherently resistant to the radiation-induced transient faults [134, 138] because the magnetization direction of the ferromagnetic (FM) layers is physically immune to energetic particles, i.e., cannot switch as a result of particle strikes. This is because the data bit is stored in the MTJs based on the spin direction, such as spin-up and spin-down (spin property of electrons), instead of electrical positive and negative charges as in the conventional CMOS circuits [13]. On the other hand, the issue of radiation-induced transient effects, i.e., Single-Event Transient (SET) and Single-Event Upset (SEU), in the CMOS-based circuits for read/write access operations remains unaddressed [139]. In addition, as the spacing between the sensitive nodes is reduced with technology size scaling, the vulnerability of SEU-immune flip-flops increases [69]. Thus, soft errors in flip-flops have become a concern due to the effect of SET induced Double Node Upset (DNU) simultaneously between the adjacent vulnerable nodes [94, 95]. Consequently, efficient solutions to completely and efficiently address this issue are sought to improve the reliability of hybrid STT-MTJs/CMOS nonvolatile sub-micron electronics [134], thereby making this technology feasible. In response to these aforementioned challenges, a power-efficient with superior soft-error resilience Non-Volatile Flip-Flop (NVFF) design is developed. It can significantly reduce the potential Single-Event Double Node Upset (SEDNU) rate due to the space reducing between the susceptible nodes. The experimental results validate its resilience, in terms of DNU-tolerance, and its efficiency, in terms of occupied area, speed performance, and power dissipation, compared to the prior work in [140] and [141]. The proposed design is the first NVFF circuit that has been demonstrated to tolerate DNU simultaneously.

5.2 Fundamentals of Magnetic Tunnel Junctions

An MTJ consists of two ferromagnetic (FM) layers, called fixed layer and free layer, which are separated by a thin oxide barrier, i.e., MgO [142]. There are two different magnetization configurations for FM layers, parallel (P) and antiparallel (AP), according to which MTJ resistance changes between R_P and R_{AP}, respectively. MTJ resistance is determined by the angle (θ) between the magnetization orientations of fixed layer and free layer due to the tunnel magnetoresistance (TMR) effect [143]. The MTJ resistance in P (θ=0°), and AP (θ=180°) states is expressed by the following equations [135]:

$$R(\theta) = 2R_{MTJ} \times \frac{1 + TMR}{2 + TMR + TMR \times \cos\theta} \tag{5.1}$$

$$R_{MTJ} = \frac{t_{ox}}{Factor \times Area \times \sqrt{\varphi}}\, exp(1.025 \times t_{ox} \times \sqrt{\varphi}) \tag{5.2}$$

$$TMR = TMR(0)/1 + (\frac{V_b}{V_h})^2 \tag{5.3}$$

Where V_b is the bias voltage, and $V_h = 0.5V$ is the bias voltage when TMR is half of the TMR_0, t_{ox} is the oxide thickness of MTJ, *Factor* is obtained from the resistance-area product value of the MTJ that relies on the material composition of its layers, *Area* is the surface area of the MTJ, and φ is the oxide layer energy barrier height.

A promising approach, which is known as spin transfer torque (STT), was proposed in [144] to switch the MTJ states. In the STT approach, a bidirectional spin-polarized current (I_{MTJ}) is required for switching the free layer nanomagnet configuration of the MTJ, as shown in Figure 5.1.

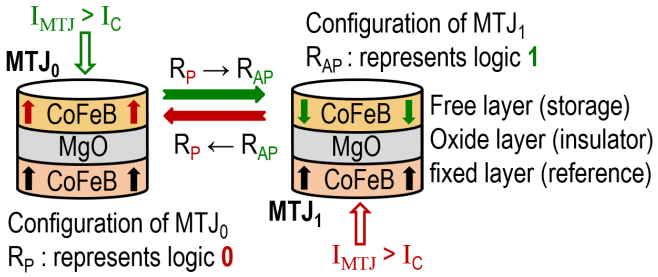

Figure 5.1: Spin transfer torque switching mechanism.

The P or AP configuration of the MTJ is determined by the direction of the current that flows through it. The required bidirectional current could be produced by means of simple MOS-based circuits. According to the relative amplitude of the I_{MTJ} and the switching critical current (I_C), STT switching behavior can be categorized into precessional region ($I_{MTJ} > I_C$), and thermal activation region ($I_{MTJ} < I_C$). To have high switching speed, MTJ is required to work in precessional region. The below equation describes the switching duration of the MTJ in this region [145]:

$$\frac{1}{\langle \tau_{STT} \rangle} = [\frac{2}{C + ln(\pi^2 \Delta)}] \frac{\mu_B P}{em(1+P^2)} (I_{MTJ} - I_C) \tag{5.4}$$

Where τ_{STT} is the mean duration for precessional switching region, $C = 0.577$ is the Euler's constant, Δ is the thermal stability factor, and m is the free layer magnetic moment. In this paper, Verilog-A is utilized to model the behavior of STT-MTJs based on the aforementioned physical equations. Then, the model is leveraged in SPICE circuit simulator to validate the functionality of the designed circuits using experimental parameters listed in Table 5.1.

Table 5.1: Parameters of STT-MTJ devices.

Parameter	Description	Value
MTJ_{Area}	$MTJ_{Length} \times MTJ_{Width} \times \frac{\pi}{4}$	40nm $\times$ 40nm $\times \frac{\pi}{4}$
t_{ox}	Thickness of oxide barrier	0.85nm
P	Spin Polarization	0.52
R_P, R_{AP}	MTJ Resistances	3.22 KΩ, 6.44 KΩ
TMR_0	TMR ratio	100%

5.3 Technology Scaling Trends on Data Bit Upset

In recent sub-micron CMOS electronics, the latching circuits have become more susceptible to Multiple Node Upset (MNU) due to the effect of SET induced DNU simultaneously. This is because the charge/discharge node capacitance of a circuit and the spacing between the vulnerable nodes are reduced to ten(s) of nanometers under the aggressive scaling of technology generations [95]. As a result, a higher Soft Error Rate (SER) is anticipated due to single particle strikes upsetting double nodes concurrently. Moreover, separating the bits of the same word becomes a complicated process, and it might not be feasible for deeply-scaled register sets, as it impacts area and power consumption [77]. Therefore, optimized solutions which offer multiple error correction capability while minimizing performance degradation, area overhead, and power dissipation, imposed by the extra logic for protection, are sought.

5.4 Radiation-Induced Transient Faults Emulation

At previous technology nodes having reduced device scaling, a primary contributor of radiation-induced soft errors was indirect ionization due to incident neutrons. However, as the amount of critical charge, Q_{crit}, becomes smaller with the process scaling, alpha particles contribute as

much as neutrons to the overall SER [28, 39]. In particular, Linear Energy Transfer (LET) from alpha particles can generate greater than 1 MeV (equivalent to 44.5fc) by direct ionization, which is more than adequate for recent technologies to charge/discharge a vulnerable node capacitance and flip its logic state [28]. It has also been reported that SEU susceptibility to low-energy alpha particles from radioactive impurities can deposit adequate charge by direct ionization. Therein, the impurities induce SEUs via their intrinsic reduction of the deposited charge threshold [16, 28]. As discussed in in Section 5.1, the MTJs are inherently immune to radiation-induced transient faults. However, the CMOS-based portion for read/write access operations needs to be hardened. To cover these cases, an accepted approach is to model the radiation-induced transient faults using a double exponential current source, which injects current to the nodes that are vulnerable to soft errors. This current source was injected/connected into vulnerable nodes during the sense mode for the master-latch and the latching mode for the slave latch. The injected current is given by [146]:

$$I_{inj}(i) = \frac{Q_{inj}}{\tau_1 - \tau_2}(e^{-t/\tau_1} - e^{-t/\tau_2}) \tag{5.5}$$

Where the parameter Q_{inj} represents the amount of injected charge which depends on the technology node size (150 fC was considered for 45nm), whereas τ_1 and τ_2 represent the time constant with typical values of 150 and 50 ps, respectively [147]. Thus, herein, α-particles effect is considered the major source of radiation-induced transient faults. As stated in [147], a transient fault strikes at an input node of a latch circuit is not a major concern since it has a marginal impact. Thus, herein, the resilience of soft fault coverage masking for the proposed DNU-immune NVFF circuit has been evaluated by considering transient faults hitting the internal (feedback loop nodes) and output nodes when the CLK single is active. Therefore, a transient pulse overlapping with the window of vulnerability can be latched and eventually cause an upset. The window of vulnerability of node i refers to the time interval during which an energetic particle striking a susceptible node i might result in a soft error [25]. Next, the developed SE-resilient NVFF design is discussed.

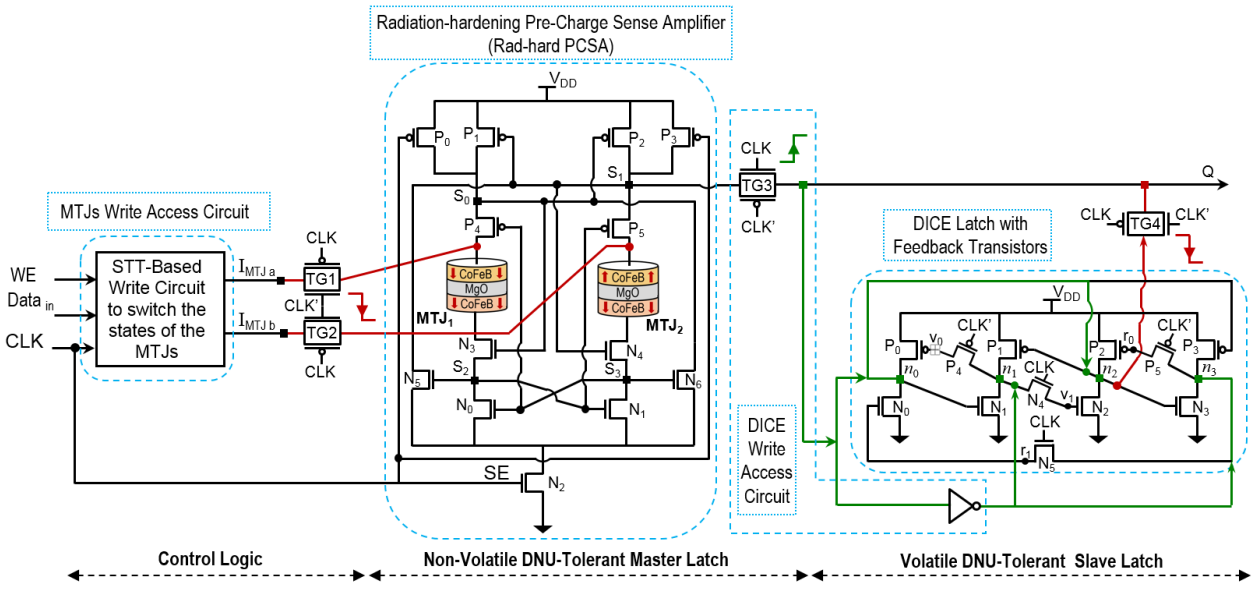

Figure 5.2: The proposed DNU-tolerant soft-error resilient nonvolatile flip-flop circuit.

5.5 Proposed DNU-Tolerant NV Flip-Flop Circuit

Herein, a power-efficient with superior soft-error resilience NVFF design is introduced. The developed design has been motivated by the techniques presented in [134] and [77], based on the combination of reliable Pre-Charge Sense Amplifier (PCSA) circuit [134] and a double upset tolerant DICE latch with feedback transistors [77]. The proposed modifications have been designed by evaluating the limitations of power consumption, soft error resilience, and volatility of CMOS-based large-scale logic circuits, regarding area overhead and speed performance degradation. Figure 5.2 depicts the design architecture of the developed NVFF circuit. The structure of the presented NVFF circuit is composed of three components including: logic control for write access operations, nonvolatile master latch, and volatile slave latch. The write circuit employs STT switching mechanism to switch the configurations of the MTJs since the STT technique is considered the most mature and advanced writing mechanism due to its low current usage, e.i., $50\mu Amps$@40nm [148], and high-speed performance. In addition, the transmission gates (TG1

and TG2) were added to isolate the write circuit from the sensing path during the sensing mode, synchronize the read and write access operations. The magnetic master latch utilizes 17 transistors in sense amplifier circuit configuration that is highly robust against particle strike induced DNU, whereas the slave latch uses 18 transistors in DICE latch with feedback transistors configuration that is substantially immune to single-event induced double upset.

5.5.1 Functionality Analysis of the Proposed NV Flip-Flop

Default operation mode: Considering the SEU-free case, when the clock signal (CLK) is low and the complement clock signal (CLḰ) is high, i.e., falling edge of the clock, the coming data bit at the input is written into the NV master latch by reconfiguring the state of the MTJs based on the bit value. This is achieved by the STT switching mechanism that generates bidirectional current to switch the free layer of the storage MTJ cell to represent a 0 or 1 logic value as compared to the reference MTJ. During the rising edge of the CLK signal, when CLK is high and CLḰ is low, the magnetic master latch is functioning in the transparent mode of operation, and the sensed data bit from the NV MTJ cells is sent to the output through TG3. Whereas when the CLK signal is low and CLḰ is high, the slave latch is functioning in the latching mode of operation, and the output is taken from n2 of DICE latch through TG4 with the falling edge of the CLK. Meanwhile, since the feedback path delay of the proposed NVFF is isolated from the input-to-output delay path during the rising edge of the clock signal, the proposed DNU-tolerant NVFF achieves excellent sensing and CLK-to-Q delay times as the performance degradation is roughly 25% compared to conventional CMOS DICE-based flip-flop, presented in [77]. The proposed NVFF was evaluated by simulations using HSPICE. Figure 5.3 shows the timing waveforms of the developed NV latching circuit. As can be seen, both nodes S_0 and S_1 are precharged to V_{DD} during the falling edge of the CLK, while the data bit and its complementary are sensed at these nodes during the rising edge of the CLK. The output, Q, is updated from node S_1 with the rising

edge of the CLK, while the logic value of node n_2 is sent to Q during the falling edge.

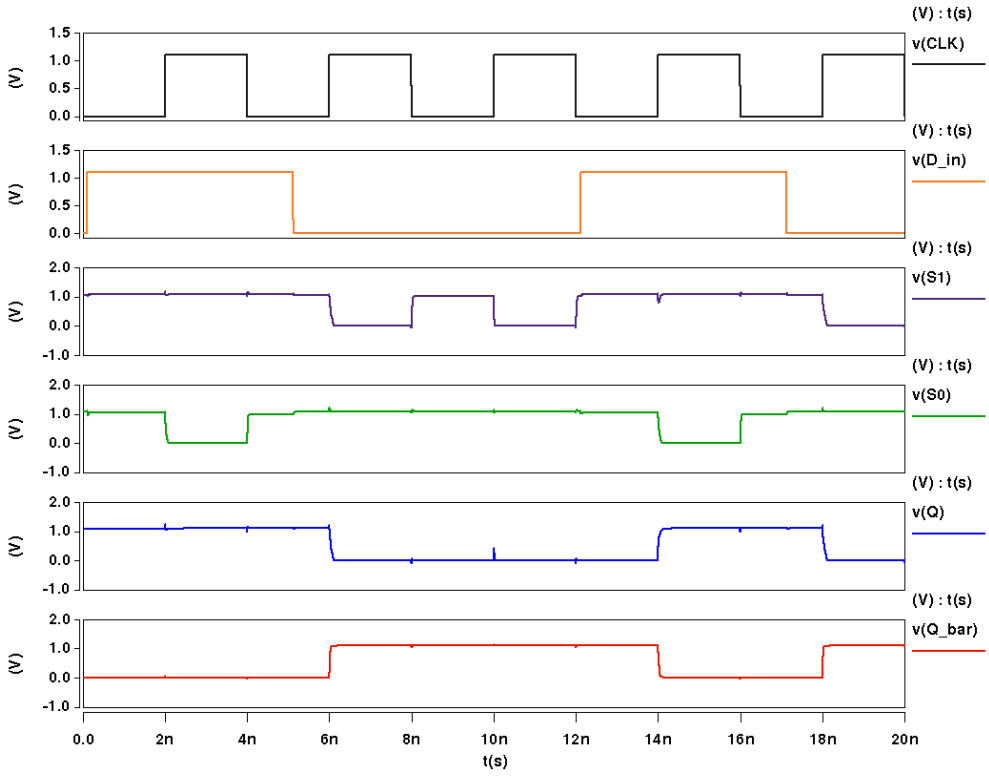

Figure 5.3: Timing waveforms of the proposed DNU-tolerant NVFF.

5.6 Experiments and Results

Experiments are carried out to analyze the overheads for the proposed soft-error tolerant NVFF design. The developed DNU-tolerant NV flip-flop design is synthesized and simulated using HSPICE simulation based on 45nm HK-MG bulk CMOS PTM-based NanGate open source library. The conducted simulations for the proposed NVFF circuit is performed with supply voltage of 1.1 V at room temperature. In order to make the design work using a single clock rate for both the TGs and the MTJs' sensing circuit, a relatively higher switching current of roughly 100 μA was used to quickly switch the configuration of the MTJs at a delay of $\sim$2 ns during the falling edge of the

clock. Thus, the frequency is set at 250 MHz for the entire circuit. However, this is at the expense of high write energy as will be discussed in Section 6. For a fair comparison purpose, all designs in [140, 141], and [77] have been re-simulated using the same CAD tools, i.e., the same technology node (45nm) and the same STT-MTJ model based on Verilog-A.

5.6.1 Evaluation of Area, Power, and Delay Overheads

Performance penalties of the proposed radiation-hardening NVFF design is analyzed and evaluated in terms of area overhead, efficiency of power consumption, and propagation delay to validate the superiority of the presented NVFF over the selected prior work in [141] and [140]. Table 5.2 summarizes the performance penalties and soft error robustness of the presented NV latching circuit, normalized to the CMOS-based flip-flop design presented in [77]. The occupied area of the proposed NVFF is estimated in equivalent device count required to construct the NVFF design, adopted from [149]. Simulation results reveal that the proposed NVFF incurs 17% and 13% of area overhead and power consumption, respectively, while the design recently presented in [140] incurs 37% of area overhead and 24% of power consumption penalty. Additionally, the proposed NVFF realizes an acceptable performance degradation (25%) in term of overall time delay (max $\{T_{CLK-Q}$ + T_{setup}, $T_{sense}\}$) and relatively associated with low leakage power (21% less than CMOS-based design). The design presented in [141] incurs the lowest area and power overheads, however, it does not provide protection for both SEU and DNU.

On the other hand, the drawback for the developed NVFF is that it consumes high energy for write access operations since a high switching current ($I_{MTJ} = 100\mu$A) with a shot duration ($\sim$2 ns) is utilized to change the states of the MTJs, but it is less than those in [140] and [141]. Thus, improved developments to lower the write energy and improve the reliability, scholastic behaviors, of STT switching mechanism is sought so that hybrid STT-MTJ/CMOS circuits can be directly

implemented on commercial technologies. The presented NVFF incurs moderate area overhead, low leakage power, and high computing performance. Thus, it can be utilized as a base component to design more complicated hybrid Spintronics/CMOS non-volatile logic and memory circuits.

Table 5.2: Performance penalties of the proposed NVFF circuit w.r.t. conventional CMOS-based flip-flop.

Metrics	**CMOS-Based FF [77]**	**NVFF [141]**	**NVFF [140]**	**Proposed NVFF**
Area (Device Count)	1	$0.57x$	$1.37x$	$1.17x$
Power Consump. (μW)	1	$0.48x$	$1.24x$	$1.13x$
Overall Delay (ns)	1	$1.72x$	$1.78x$	$1.25x$
Leakage Power (μW)	1	$0.43x$	$0.91x$	$0.79x$
Write Energy (fJ)	1	$10 \sim 11x$	$9x$	$5 \sim 6x$
Nonvolatility	x	✓	✓	✓
SEU / DNU Tolerant	✓/✓	x/x	✓/x	✓/✓

5.6.2 Evaluation of SEU and DNU Immunity

Table 5.3 lists the total number of vulnerable nodes that cause soft errors. In addition, the most vulnerable nodes of each latch (master/slave) are identified based on their susceptibility. These nodes, i.e., (S0, S1, S2, and S3) for master-latch and (n0, n1, n2, n3, v0, and v1) for slave-latch, result in a latch upset when hitting by an energetic particle with adequate LET to flip the latch state. These nodes are considered susceptible area because they are connected to the reverse-biased junction of transistors in OFF state. In order to quantify the robustness of the presented design against radiation-induced soft errors, a transient current source was connected to each susceptible node,

identified in Table 5.3, with a maximum pulse amplitude of 0.192 mA and 800 ps of duration. These HSPICE simulations were carried out to determine the vulnerable node pairs, i.e., most susceptible to particle hits, so that if they suffered a hit and flipped the logic state, then the selected pair is considered unprotected. This current source was injected into a single node to emulate a SEU, whereas double current sources were connected to double adjacent susceptible nodes, node pair, concurrently to imitate the effect of DNUs. The results of the transient simulation demonstrate that the presented NVFF is robust to achieve 100% of SEU masking, while the probability of unmasked DNU is found to be roughly 3.3%, as listed in the last column of Table 5.3. The vulnerability of DNU results from the fact that the hardened PCSA circuit has two vulnerable node pairs (S0 and S3; S1 and S3), and the DICE latch has a single susceptible node pair (n0 and v0), shown in Figure 5.2. However, this vulnerability can be alleviated by isolating the device sensitive drain area occupied by these node pairs on the cell's layout, which is beyond the scope of this work.

Meanwhile, among the designs listed in Table 5.2, the presented NVFF is the only non-volatile latching circuit that can highly tolerate SET induced DNU simultaneously. It is worth noticing that each latch circuit of the proposed NVFF is able to tolerate DNU concurrently, which makes the developed NVFF circuit an attractive soft-error resilience storage element for nanoscale device technology.

Table 5.3: Soft-error tolerance analysis of the proposed NVFF design.

Latching Circuit	**Total ♯ of Nodes**	**♯ of Vulnerable Nodes**	**Possible ♯ of Node Pairs**	**Unprotected Node Pairs**	**Fault Masking Coverage (%)**
Master-Latch	6	4: (S0, S1, S2, S3)	15	(S0,S3) and (S1,S3)	100 for SEU 86.67 for DNU
Slave-Latch	8	6: (n0, n1, n2, n3, v0, v1)	28	(n0,v0)	100 for SEU 96.43 for DNU
The Proposed NVFF	14	10	91	3 pairs	100 for SEU 96.7 for DNU

5.7 Summary

In this chapter, a soft error resilient NVFF circuit that can achieve intriguing attributes is presented. The developed NVFF achieves reduced area and power penalties (14.6% and 10.3%, respectively), while offering speed performance improvement of 29.5% as compared to the prior work in [140]. In addition, it significantly reduces leakage power by 21% compared to the CMOS-based design. Thus, the proposed soft-error resilient NVFF can be utilized to harden the critical or vulnerable registers against radiation-induced transient faults, including both SEU and DNU. Meanwhile, since the STT writing mechanism incurs high energy for switching the MTJs states, thus energy-efficient NV-FF circuits are still sought. In the next chapter, energy-efficient Spin Hall Effect (SHE)-MTJ NVFF designs will be introduced as promising candidates to realize energy-aware and reliability-aware soft-error resilient latching circuits.

CHAPTER 6: ENERGY-AWARE AND RELIABILITY-AWARE NON-VOLATILE SPINTRONIC REGISTERS FOR RESILIENT ARCHITECTURAL STATE STORAGE

This chapter covers the designs for energy-aware and reliability-aware soft-error resilient latching circuits to improve soft-error masking robustness of emerging large-scale NVFF circuits while concurrently reducing its power dissipation to realize energy efficiency. In particular, low-energy radiation-hardening approaches are proposed to develop Non-Volatile (NV) Flip-Flop (FF) circuits using spintronic devices. Also, Spin Hall-Effect Magnetic Tunnel Junctions (SHE-MTJs) are used to design a radiation-hardened NV-latch that is proposed to be utilized as a shadow latch to maintain the data during the standby mode when the circuit is power-gated. Moreover, soft error resilient CMOS-based latching circuits are designed to be leveraged as master and slave latches in the NV-FF structure. The proposed hardening techniques are based on using feedback loops and clock-gating Muller C-elements, as well as increasing the charge capacity of vulnerable nodes. The circuit simulations indicate that the proposed Single-Event Upset (SEU) and Double Node Upset (DNU) resilient latching circuits can achieve at least 81% and 24% power-delay-product improvement, respectively, while incurring comparable area overhead compared to the previous energy-efficient radiation-hardened latch designs. Finally, the proposed latching circuits are combined to develop four radiation-hardened NV-FF designs. The results obtained show that, using the proposed NV latching circuit as a shadow latch can result in two orders of magnitude reduction in energy consumption compared to the FF circuits with an NV master latch. In addition, the proposed latches achieve favorable tradeoffs in terms of minimized performance overheads and maximized robustness (100%) of *Soft Fault Coverage* to single and double upset.[1]

[1] © 2019 IEEE. Reprinted, with permission, from [150].

6.1 Writing Mechanisoms for Switching the MTJs States

Spin transfer torque (STT) switching mechanism is one of the most commonly-used methods for changing the MTJ states [144]. However, the STT approach incurs high switching energy, and suffers from oxide barrier reliability issues due to the shared read and write path. Therefore, Spin Hall Effect (SHE)-assisted STT approach has been introduced in recent years, which consumes significantly less switching energy consumption while providing separate paths for read and write operations [11, 151].

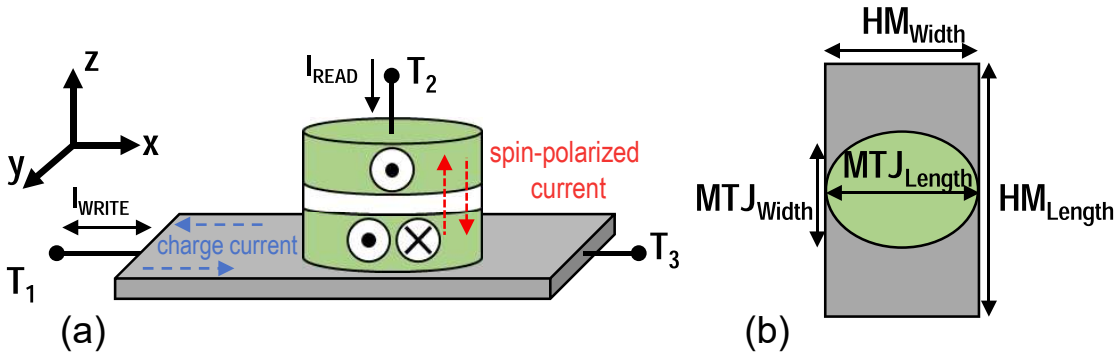

Figure 6.1: (a) SHE-MTJ vertical structure, (b) SHE-MTJ top view.

Figure 6.1 shows a three terminal SHE-MTJ device, in which the free layer is directly connected to a Heavy Metal (HM). The MTJ state can be changed via a charge current applied to the HM, which can generate a spin-polarized current that flows through the MTJ free layer. In particular, a positive current along +x induces a spin injection current in +z direction. The injected spin current produces the required spin torque for aligning the magnetic direction of the free layer in +y directions, and vice versa. The ratio of the injected spin current to the applied charge current is normally greater than one in SHE-MTJ devices, and is called Spin Hall Injection Efficiency (SHIE), which is expressed in Eq. 6.1 [11], where λ_{sf} is the spin flip length in HM and θ_{SHE} is the

spin hall angle.

$$SHIE = \frac{\pi.w_{MTJ}.l_{MTJ}}{4.t_{HM}.w_{HM}}\theta_{SHE}\left[1 - sech(\frac{t_{HM}}{\lambda_{sf}})\right] \tag{6.1}$$

The relation between the switching delay (τ_{SHE}) and the applied charge current (I_{WRITE}) is shown in Eq. 6.2, where v_c is the critical switching voltage given by Eq. 6.3 and R_{HM} is the HM resistance.

$$\tau_{SHE} = \frac{\tau_0 ln(\pi/2\theta_0)}{(R_{HM}I_{SHE}/v_c) - 1} \tag{6.2}$$

$$v_c = \frac{8\rho I_c}{\pi\theta_{SHE}l_{HM}\left[1 - sech(\frac{t_{HM}}{\lambda_{sf}})\right]} \tag{6.3}$$

$\tau_0 = qM_SV_{HM}/I_cP\mu_B$ is the characteristic time, $\theta_0 = \sqrt{k_B/2E_b}$ is the effect of stochastic variation, and E_b is the thermal barrier of the magnet. In this paper, the SHE-MTJ devices are modeled based on the modeling approach proposed in [143] and [152] using the aforementioned equations. In particular, a Verilog-AMS model is developed and utilized in SPICE circuit simulation to validate the functionality of the designed circuits. Table 6.1 lists the experimental parameters for an in-plane CoFeB/MgO/CoFeB MTJ with β-tungsten (β-W) HM that are obtained from [11, 12].

6.2 Soft Fault Coverage Analysis of Latching Circuits

Here, to evaluate the radiation tolerance of the proposed circuits, the *Soft Fault Coverage (SFC)* metric is introduced, which is determined by the ratio of unprotected node pairs over the total

number of vulnerable node pairs:

$$SFC = \left(1 - \frac{\binom{m}{2}}{\binom{n}{2}}\right) * 100\% \tag{6.4}$$

where n is the total number of sensitive nodes, and m represents the total number of unprotected nodes in a circuit.

Table 6.1: Parameters of SHE-MTJ devices [11, 12].

Parameter	**Description**	**Value**
HM_{Volume}	$\mathrm{l}_{HM}\times \mathrm{w}_{HM}\times \mathrm{t}_{HM}$	100nm $\times$60nm $\times$3nm
MTJ_{Area}	$\mathrm{l}_{MTJ}\times \mathrm{w}_{MTJ}\times \frac{\pi}{4}$	60nm $\times$30nm $\times\frac{\pi}{4}$
MTJ_{Area}	Reference MTJ Surface	50nm $\times$25nm $\times\frac{\pi}{4}$
α	Gilbert Damping factor	0.007
M_S	Saturation magnetization	$7.8e5$ A.m^{-1}
P	Spin Polarization	0.52
R_P, R_{AP}	MTJ Resistances	2.8 KΩ, 5.6 KΩ
R_P	Reference MTJ Resistance	4.1 KΩ
TMR	TMR ratio	100%
θ_{SHE}	Spin Hall Angle	0.3
ρ_{HM}	HM Resistivity	200 $\mu\Omega$.cm
λ_{sf}	Spin Flip Length	1.5 nm
I_{C-SHE}	SHE-MTJ Critical Current	108 μA

Transient faults are known to have negligible impacts on the inputs nodes of a latch circuit [147]. Thus, herein, focus is on faults affecting internal nodes of feedback loops and output nodes of the circuits to evaluate the radiation-tolerance of the proposed designs when the data is latched with the rising/falling edge of the clock signal. Therefore, a transient pulse overlapping with the window of vulnerability can be latched and cause an upset. The window of vulnerability refers to the period during which a particle striking a sensitive node might induce a soft error [25]. Vulnerable nodes of a circuit are identified based on their susceptibility, whereby nodes connected to the reverse-biased

junction of transistors in the OFF state are considered sensitive nodes [153].

6.3 Proposed Soft-Error Resilient NV-FF Designs

Herein, resilient spintronic NV flip-flop designs are developed to significantly reduce the impact of transient and upset glitches, thus elevating the robustness of latching circuits. This is a vital focus whereas the feedback and output nodes of a latching circuit are the most vulnerable nodes capable of contributing to an SEU [147]. The NV-FF circuits are designed by considering hardening for these susceptible nodes. This can maximize the soft error immunity while minimizing the area and energy costs of the proposed latching circuits. First, the design of the proposed spintronic latching circuit is discussed.

6.3.1 DNU-Resilient Non-Volatile Latch Circuit

The proposed radiation-hardened non-volatile latch circuit is shown in Figure 6.2, in which the write circuit has not been depicted for simplicity. Transmission Gates (TGs) are used in the SHE-MTJ write circuit, which are characterized by their near-optimal full-swing switching behavior. TGs can provide an energy-efficient and process variation resilient write circuit as established in [154]. Figure 6.3 (a) and (b) show the structure of the TG-based write circuit required for switching the logic states of NV-latch's SHE-MTJ and reference SHE-MTJ cells, respectively. As shown, if Q=V_{DD} and WE=V_{DD}, the TGs will generate a charge current (I_C) in –x direction, which induces a spin current (I_S) in +z direction aligning the magnetic direction of the free layer (FL) in –y direction. Thus, the magnetization orientation of the free layer and fixed layer will be in the AP state, leading to a high resistance for the SHE-MTJ device (R_{MTJ}=R_{HIGH}=R_{AP}). Similarly, Q=GND aligns FL in +y magnetic direction, leading to a SHE-MTJ with P state, thus R_{MTJ}=R_{LOW}=R_P.

The write circuit for the reference SHE-MTJ is only utilized after the device power-on, in which the magnetic direction of reference SHE-MTJ is tuned to the P state by applying Q_{REF}=0 and WE_{REF}=1 signals. The TG-based write circuit can generate a current amplitude of ∼160 μA, which is sufficient to switch the logic state of MTJ in less than 2ns. The maximum write current density for the SHE-MTJ device utilized in this paper is approximately 9×10^{11} A.m^{-2}, which is greater than the minimum write current density required for switching the conventional 2-terminal MTJs. However, the reliability of tunneling oxide barrier, which is a major reliability issue in 2-terminal MTJs, is improved in the SHE-MTJ, since the current does not flow through it during the write operation [155]. Moreover, the authors have investigated the effect of radiation particles on the SHE-MTJ write circuitry in [156], where it was shown that the SHE-MTJ magnetic state does not change due to the short duration of the radiation-induced current pulses. Thus, herein, concentration is on protecting the read operation of the latch in the presence of radiation particles.

The read circuitry is designed based on the well-known Pre-Charge Sense Amplifier (PCSA) circuit [134]. The PCSA compares the stored resistive value of the SHE-MTJ cells with a reference MTJ cell. As expressed by Eq. 5.1 and Eq. 5.2, the resistance of the SHE-MTJ is determined by the angle (θ) between the magnetization orientations of its fixed layer and free layer. In particular, if the magnetization orientation of the layers is in parallel (P) state ($\theta = 0$), the SHE-MTJ resistance is low (R_{LOW}=R_P), and if the magnetization orientation of the layers is in anti-parallel (AP) state ($\theta = \pi$), the SHE-MTJ resistance is high (R_{HIGH}=R_{AP}). Thus, the resistance of the reference SHE-MTJ should be designed such that its resistance value in P configuration ($R_{P-ref-MTJ}$) is between the resistances of the NV-latch's SHE-MTJ cell ($R_{Latch-MTJ}$) in P and AP states as expressed by Eq. 6.5. Specifically, the resistance of the SHE-MTJ cell can be tuned by adjusting the area of the reference MTJ cell according to Eq. 5.2.

$$R_{P-ref-MTJ} = \frac{1}{2}(R_{P-Latch\ MTJ} + R_{AP-Latch\ MTJ}) \tag{6.5}$$

where,

$$NV_{out} = \begin{cases} 1 & \text{if } R_{Latch-MTJ} > R_{P-ref-MTJ}; \\ 0 & \text{if } R_{Latch-MTJ} < R_{P-ref-MTJ}. \end{cases}$$

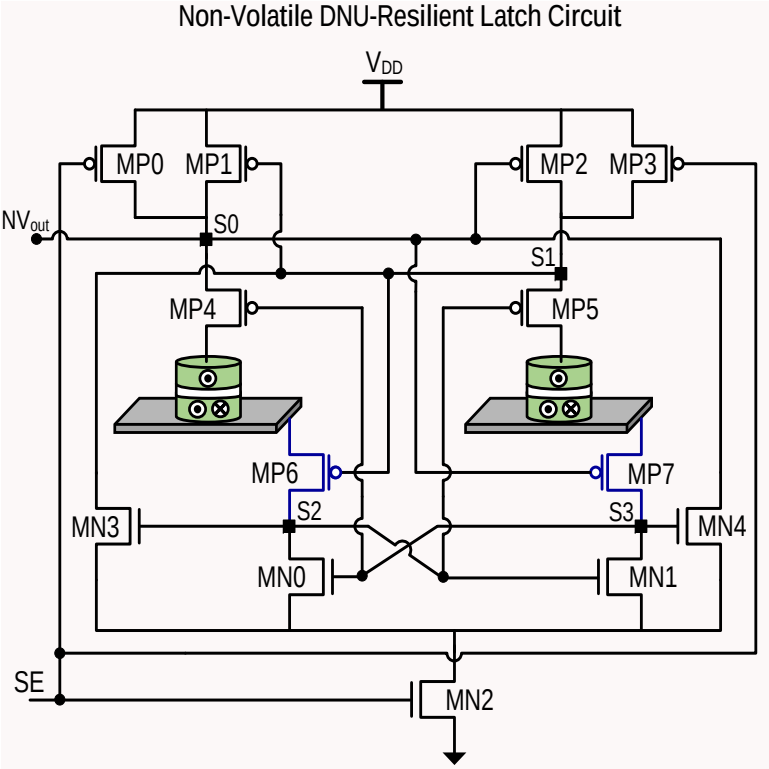

Figure 6.2: Circuit diagram for the NV DNU-resilient latch, in which the write circuits have not been depicted for simplicity (refer to Figure 6.3).

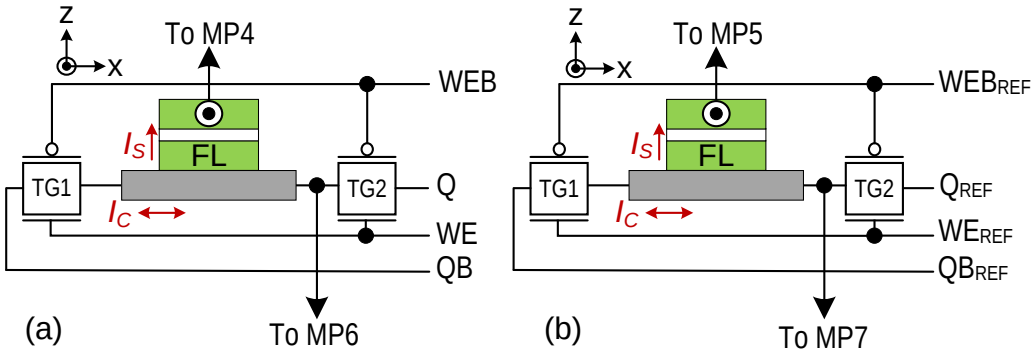

Figure 6.3: Write circuit for (a) NV-latch's MTJ cell, (b) Reference MTJ cell.

Therefore, if the resistive value of a SHE-MTJ cell in the NV-FF circuit is greater than the resistance of the reference cell, the output of the PCSA will be logic '1' and vice versa. Read operation using PCSA involves two operating phases: (1) *pre-charge*, in which S0 and S1 nodes are charged to the nominal voltage (V_{DD}), and (2) *discharge*, where SE signal equals V_{DD} and the MP0 and MP1 transistors are OFF. Therefore, the pre-charged S0 and S1 nodes are disconnected from the voltage source, and start discharging with different speeds due to the difference between resistances of the discharge paths. For instance, if the voltage of node S0 discharges faster, the MP2 transistor will be ON earlier than MP1 due to the voltage difference between its gate and source terminals. This recharges node S1 to V_{DD}, which results in MP1 transistor remaining OFF and node S0 completely discharging to the ground. In the proposed circuit, node S0 is connected to the storage element and shows the output of the latch, while node S1 is connected to a reference SHE-MTJ cell. The dimensions of reference MTJ is designed such that its resistance in *P* configuration is between the resistance of the storage cells in *P* and *AP* states.

The internal node S0 and S1 are the radiation-sensitive nodes of the proposed NV latch circuit. The hardening circuitry includes four NMOS (MN0-MN4) and two PMOS (MP6-MP7) transistors that are used to discharge the radiation-induced electric charges injected to these vulnerable nodes (S0 and S1). However, the utilization of these feedback transistors creates two sensitive nodes within the hardening circuitry, S2 and S3, which are the surroundings of the reverse-biased drain junction of MP6 and MP7 transistors during the discharge phase. The sensitivity of these nodes depends on the state of the SHE-MTJ. In particular, when the SHE-MTJ is in *P* configuration, then S3 is the sensitive node. While node S2 is the vulnerable node when SHE-MTJ cell is in *AP* state. To increase the radiation tolerance of S2 and S3 nodes, the MP6 and MP7 transistors connected to these vulnerable nodes are enlarged to increase their equivalent charge capacitance.

Soft-error resiliency of the proposed NV-latch is evaluated via SPICE circuit simulations, in which the latch storage cell is considered in *P* state, thus S0, S1, and S3 are the sensitive nodes. Figure 6.4

shows the behavior of the proposed radiation-hardened NV-latch circuit in presence of SEUs. An SEU on node S1 increases its voltage, however since the voltage of S0 is still near zero, MP2 remains ON and S1 will be re-charged to V_{DD} through MP2. An SEU on node S0 increases its voltage level to more than V_{DD}, however since the voltage of S3 is near V_{DD}, the MN4 transistor remains ON and the injected charge will be discharged through MN4 and the output will be recovered after ~ 0.9 ns. Radiation particles striking at node S3 have a negligible effect on its voltage, thereby the voltages of the nodes S0 and S1 remain unchanged. The transient behavior of the proposed NV-latch in presence of DNUs is shown in Figure 6.5. Particles striking S0 and S1 nodes do not impact the voltage level of node S3, therefore MN4 will remain ON and the injected charge at node S0 will be discharged through MN4 leading to the S1 being recharged to V_{DD} through MP2. DNUs on nodes S3 and S0 have a minor impact on the voltage of S3 due to its increased charge capacitance, thus MN4 remains ON and the injected charge will be discharged through MN4 and the output will be recovered. Similarly, DNUs on node S3 and S1 result in an insignificant effect on S3 since the voltage of S0 remains near zero, and S1 will be charged to V_{DD} through MP2.

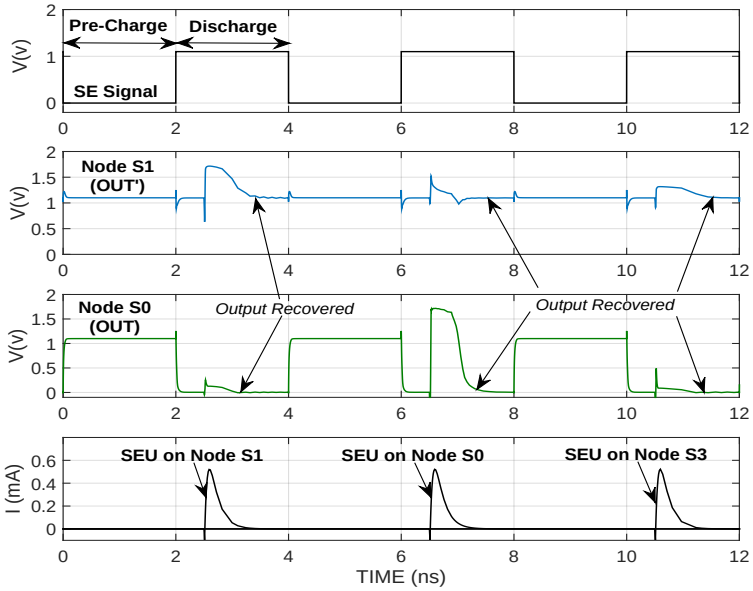

Figure 6.4: Transient response of the proposed NV-latch to SEUs.

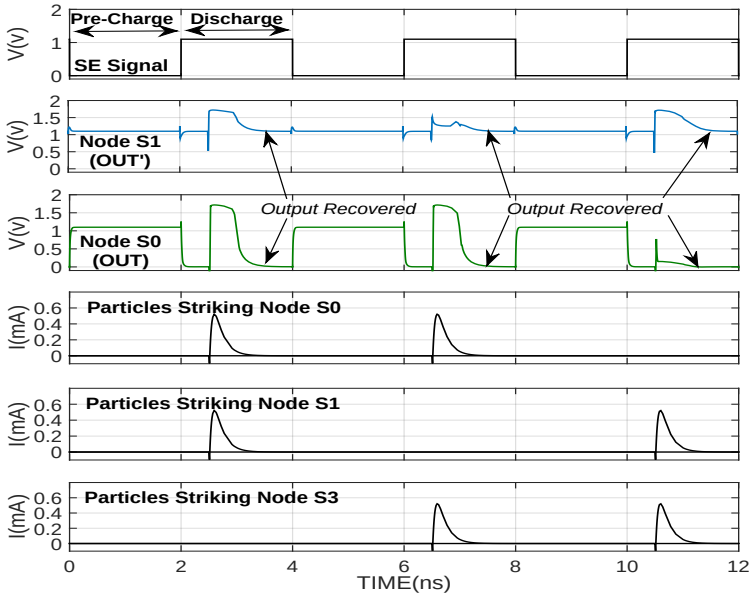

Figure 6.5: Transient response of the proposed NV-latch to DNUs.

Herein, the proposed SEU and DNU resilient CMOS-based latch circuits can be used as master and slave latches in the NV-FF designs, as shown in Fig. 6.6 and 6.8, respectively. The proposed DNU-resilient NV-latch is utilized to backup data only during standby mode. Hence, the logic state of the CMOS-based flip-flop is stored into the MTJs before entering standby mode via power gating circuits. In the wake-up mode, the data stored in MTJ is sensed and transferred to the CMOS-based logic. In this case, the MTJ cells are only written in the standby mode, while CMOS-based latches are used during the normal operation.

6.3.2 SEU-Resilient Non-Volatile Flip-Flop Circuit

Herein, the MTJs serve to backup data only during standby mode, whereas the proposed SEU-resilient NV-FF operates in four modes including: *(1) latch mode:* this is the normal operation of the circuit which is similar to the conventional FF circuits, *(2) back-up mode:* the data in slave latch is stored into the NV-latch, *(3) standby mode:* the circuit is power-gated and the data is maintained

in non-volatile elements, i.e. MTJs, with near-zero leakage power, *(4) wake-up mode:* the stored data is sensed using proposed radiation-hardened PCSA and recovered for normal latch operation.

The proposed SEU-resilient CMOS-based latching circuit shown in Fig. 6.6 includes an interlocked loop, a regular latch, and a clock-gating output comparator (C-element) circuit. Transmission gates are used to separate master and slave latches, and to isolate the CMOS-based datapath from the NV-latch circuit. During the restore mode, the data from NV-latch is sensed and interlocked into the feedback loop of slave latch. During the latching phase, the proposed circuit can completely mask transient faults impacting a single vulnerable node that is located either in the datapath logic or in the feedback loops. To reduce the area overhead and power consumption of the hardening circuitry, the inverters required for each redundant interlocked loop are replaced with a single C-element circuit that performs both comparison and inversion operations on the data bit, therefore a masked output is generated at the internal nodes. The functionality of the proposed SEU-resilient latch circuit used as a master latch is described below. Similar circuitry is utilized as the slave latch. When the clock signal (CLK) is low, the input data bit is propagated to the output of the master-latch through TG1, and simultaneously stored into the regular latch configured by INV1 and INV2 through TG2. The datapath delay (CLK-to-Q) equals the propagation delay of a single TG realizing high speed operation. In the latch mode, when the CLK signal is high and its complementary (CLKB) is low, first the feedback from output node and regular latch (n1) are compared through the C-element circuit (configured by MP0-MP1 and MN0-MN1), which forms an interlocked loop and produces the masked internal output, i.e. node n4. Then, the logic value of node n4 is compared with the output of the regular latch (n2) via a clock-gating C-element circuit (constructed by MP2-MP4 and MN2-MN4) to produce the masked output of the latch. Thus, a radiation-induced transient pulse striking node n1 or n3 is filtered out by the internal interlocked comparator, while the output comparator masks the transient faults striking at node n2 or n4. In addition, during the normal operation the SE signal in the feedback loops set to logic '0', thus the

latched data bit in the feedback loops will be updated from the coming data at the input. During the wake-up mode, the stored logic bit in the NV logic is sensed and will be transferred to and latched in the feedback loops of the slave-latch via TG8 and TG9. In this case, TG8 and TG9 become ON, thus the feedback loops will be updated from the NV-latch circuit before resuming normal operation. Therefore, the SE signal is essential to control and synchronize the feedback loops write operation of the slave-latch.

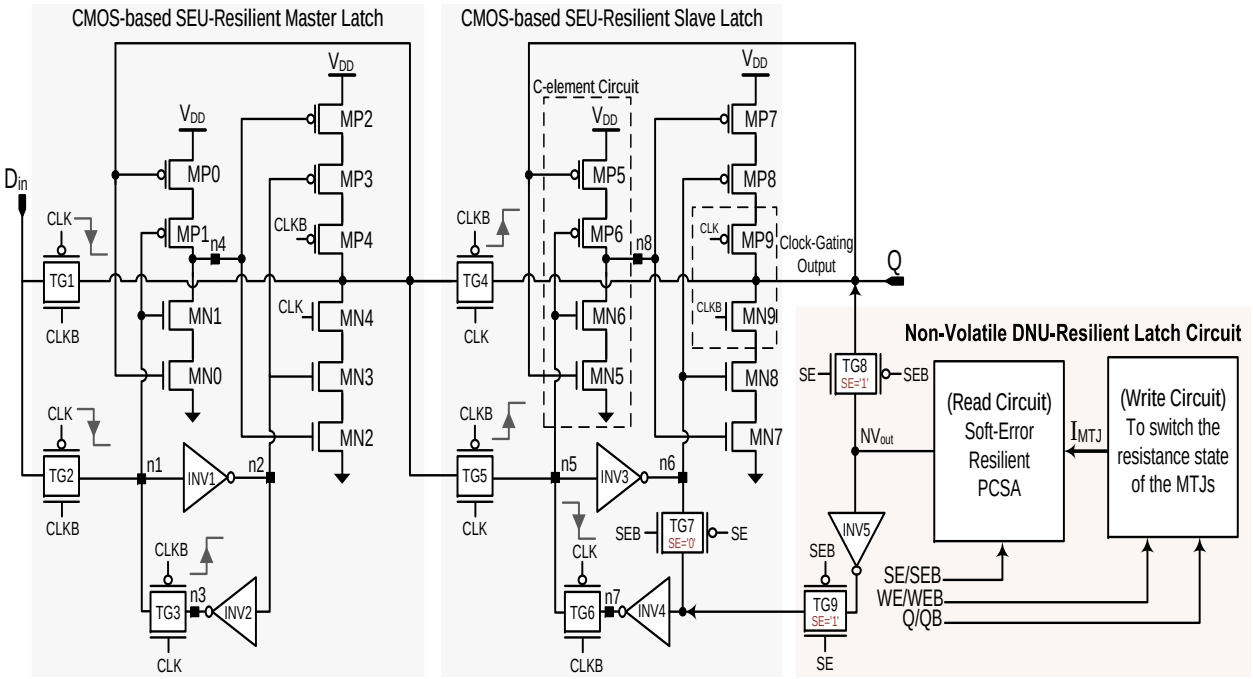

Figure 6.6: Circuit diagram of the proposed high-performance shadow-based SEU-resilient NV flip-flop.

The radiation-tolerance of the proposed SEU-resilient NV-FF was evaluated by SPICE circuit simulation. Radiation-induced current pulses were injected to sensitive nodes n1, n2, n3, and n4 at different instances of time, as shown in Fig. 6.7. The logic states are recovered in n1, n2 and n4 nodes after $\sim$ 200 ps, while node n3 temporarily holds an incorrect logic value until the next rising

edge of the CLK. Meanwhile, the output (Q) remains unchanged, and the proposed SEU-resilient latch maintains correct operation. It is worth noting that transient faults striking the output of TG1 do not impact the latch output since it is in transfer mode and its data is invalid during this period of time [147]. Although the proposed SEU-resilient NV-FF achieves complete SEU masking, it can not tolerate multiple upsets striking sensitive nodes in master or slave latches simultaneously.

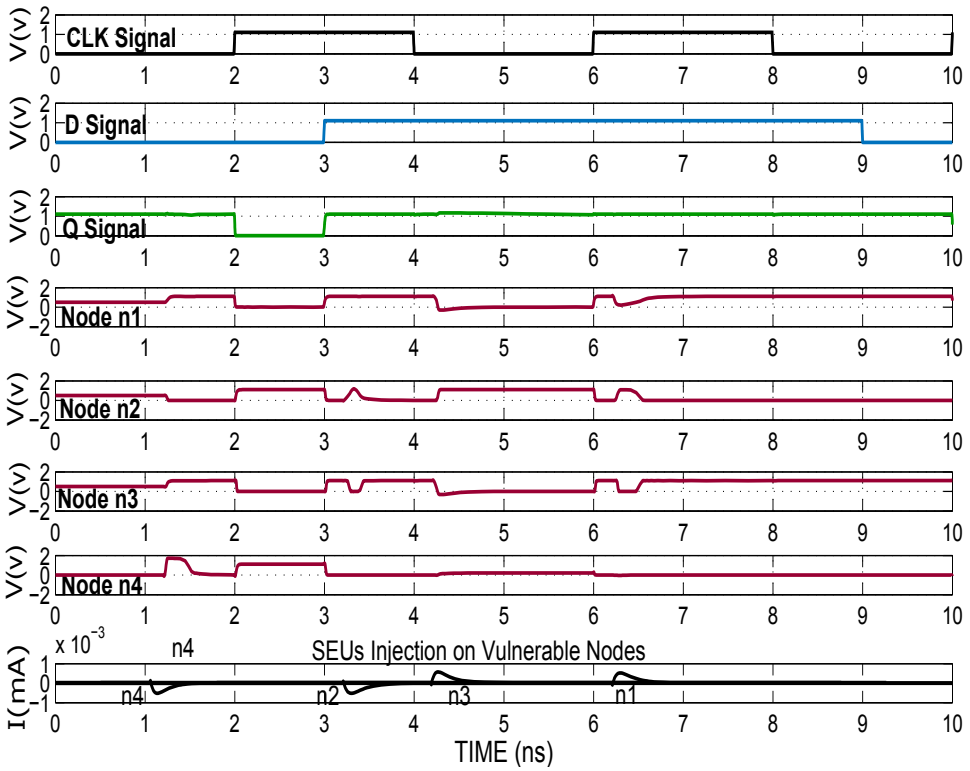

Figure 6.7: Simulation waveforms for the proposed SEU-resilient latch when a transient fault strikes a sensitive node.

6.3.3 DNU-Resilient Non-Volatile Flip-Flop Circuit

Herein, a high-performance and DNU-Resilient NVFF (DNUR-NVFF) circuit is proposed, which realizes radiation-tolerance by triplicating/duplicating the susceptible nodes, and using an output feedback loop to construct protected internal interlocked nodes. In addition, to protect the output

node, duplicated interceptive feedback loops are activated during the latching phase. Without loss of generality, the latching phase is assumed to occur during the rising edge of CLK. In this regard, the output changes according to the majority logic values of internal nodes.

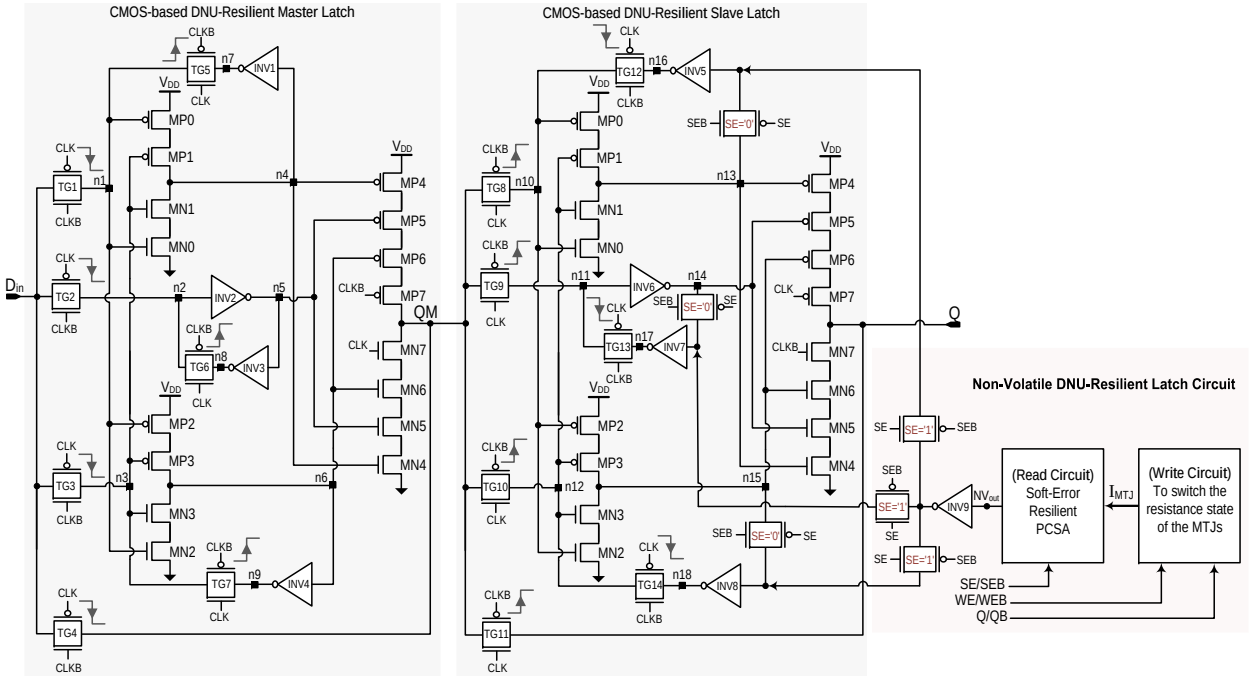

Figure 6.8: Circuit diagram of the proposed high-performance shadow-based DNU-resilient NV flip-flop.

Figure 6.8 depicts the structure of the proposed DNUR-NVFF. Similar to the SEU-resilient NV-FF design, each latch consists of a dual redundant interlocked loop, a regular latch, and clock-gating output comparators. Moreover, TG4 is added to transfer data to the output (Q) during the transparent mode of operation. This reduces the propagation delay time (CLK-to-Q) and achieves high speed operation. Herein, a C-element circuit is employed in the logic datapath of the latches to protect the sensitive nodes while reducing the area overhead and power consumption of the redundant logic. The functionality of the proposed DNU-resilient latch circuit is described below. Next, the focus is on the behavior of the slave latch since a similar circuit is also utilized as master

latch.

During the transfer mode, when the clock signal (CLK) is high, the input data bit is transferred to the output (Q) through TG11, and simultaneously stored to the regular latch (configured by INV6 and INV7) through TG9. Likewise, it is locked into the upper interlocked feedback loop (constructed by MP0-MP1 and MN0-MN1) and lower loop (configured by MP2-MP3 and MN2-MN3) through TG8 and TG10, respectively. Thus, the datapath delay (CLK-to-Q) is equal to the propagation delay of TG11, thereby realizing a significantly high speed operation. In the latching phase, the CLK signal is low and the behavior of the DNUR latch is similar to that of the proposed SEU-resilient latch, except that the output of the lower loop (n15) is now driven by a C-element (composed by the MP2-MP3 and MN2-MN3), which prevents transient faults striking n18 node from being propagated to node n15. In addition, during the latching phase the feedback from the upper and lower loop is compared to ensure the integrity of data bit at the internal nodes. Thereafter, the logic state of n13, n14, and n15 nodes are compared via a clock-gating C-element circuit (constructed by MP4-MP7 and MN4-MN7) to produce the masked output of the latch. Transient faults striking nodes n16 and n18 temporarily alter their logic state, however the correct logic state will be recovered via a C-element circuit and the output data bit will remain unchanged. Hence, the integrity of nodes n10, n12, n16, and n18 is ensured via the upper and lower comparison (C-element) circuit, while nodes n13, n14, and n15 are protected via the output comparator. Therefore, radiation-induced upsets striking single or double sensitive node(s) will be masked through the internal and output comparators. Here, the SE signal is performing the same as in the SEU-resilient NV-FF design.

To investigate the radiation-tolerance of the proposed DNUR-NVFF design, SPICE circuit simulations were performed in presence of DNUs, as shown in Fig. 6.9. To model the effect of DNUs, radiation-induced current pulses are simultaneously applied to the susceptible nodes of the design. The injected upsets temporarily alter the logic state of the struck nodes, however the impacted

nodes n3, n4, n5, and n6 are quickly recovered through the interlocked feedback loops, and the final output (Q) is not affected. Moreover, transient faults striking the output of TG11 will not impact the output of the latch since it is in transparent mode and its data is not valid in this mode. Hence, the proposed DNUR-NVFF design is able to produce the correct output, Q, in presence of DNUs striking master latch, slave latch, and/or the NV latching circuit, while realizing reduced area overhead, delay, and power consumption, as discussed in the next section.

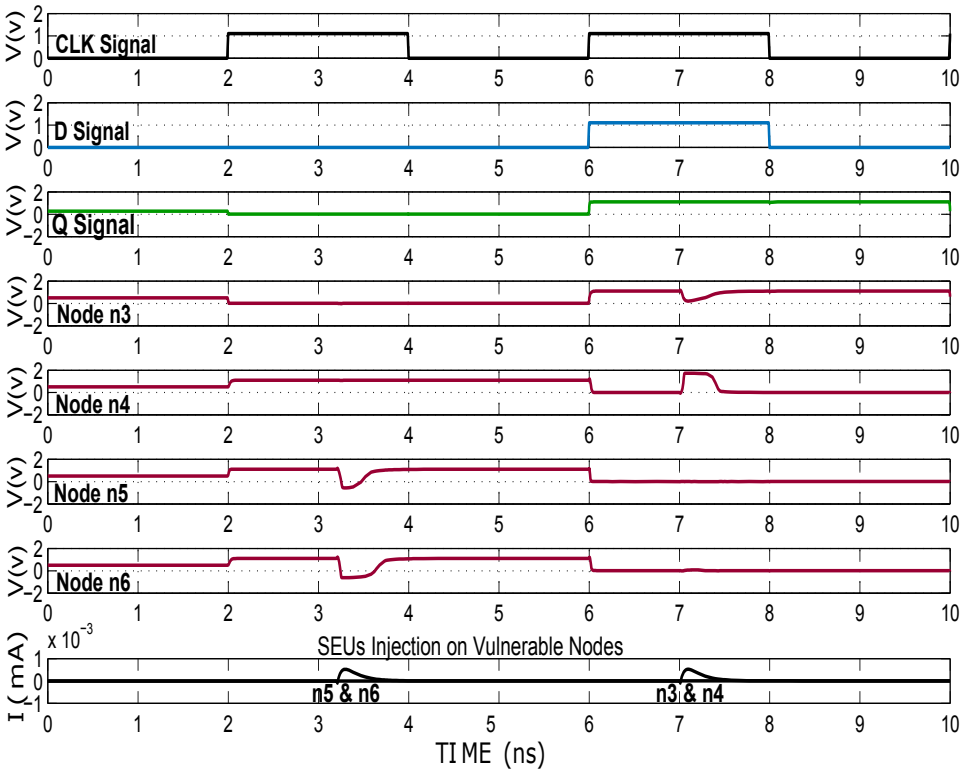

Figure 6.9: Simulation waveforms for the proposed DNU-resilient latch design when transient faults hit double vulnerable nodes.

6.4 Results and Discussion

To evaluate the performance of the aforementioned SEU and DNU resilient NV-FF designs, the proposed circuits are implemented by SPICE circuit simulator using 45nm CMOS technology [131] and the SHE-MTJ model introduced in section 6 with 1.1V nominal voltage at room temperature. To provide a fair comparison, the designs previously presented in [58, 68, 77, 157] have also been implemented using the same 45nm CMOS library.

6.4.1 Performance Analysis of Proposed Latching Circuits

Table 6.2 provides a comparison between the proposed SEU and DNU tolerant latching circuits and the soft-error resilient designs recently introduced in [58, 68, 77, 157]. The last four columns of the table list the relative overheads in terms of area (ΔArea), power consumption (ΔPower), datapath delay (ΔDelay), and PDP (ΔPDP) compared to the conventional TMR-based latch [158] as baseline.

Simulation results show that the proposed SEU-resilient latch realizes 70.45% and 92.81% reduction in power consumption and delay (CLK-to-Q) compared to the baseline design, respectively. Moreover, the presented SEU-resilient design achieves 97.87% decrease in terms of PDP, which is the largest PDP improvement compared to the previous energy-efficient SEU-tolerant latches [77, 157]. Additionally, the proposed SEU-resilient latch circuit uses only 20 MOS transistors, which makes it the second most area-efficient SEU-tolerant design after [77] with 18 transistors. On the other hand, the last three rows of Table 6.2 provide a comparison between the proposed DNU-resilient latch circuit and prior work recently presented in [58] and [68]. The results obtained indicate that the proposed DNU-resilient latch reduces area and power consumption by 29.63% and 27.49%, respectively, while increasing the speed by 90.63% compared to the base-

line design. Moreover, as listed in the last column of the table, the DNU-resilient latch circuit realizes 93.21% PDP improvement compared to the baseline design, making it the most energy-efficient design among the DNU-resilient latching circuits investigated herein. Thus, the comparison results verified the advantages of the proposed SEU and DNU resilient latch circuit in terms of PDP, which is mainly achieved by significantly reducing the propagation delay while maintaining power-consumption and area overhead values comparable to previous high performance radiation-hardened latch designs.

Table 6.2: Comparison of the proposed soft-error resilient latches and previously presented radiation-hardened designs.

Design	Area (Device Count)	Delay (ps)	Power (μW)	PDP	SEU Tolerant	DNU Tolerant	ΔArea (%)	ΔDelay (%)	ΔPower (%)	ΔPDP (%)
Baseline [158]	54	64.72	64.2	4155.02	✓	×	-	-	-	-
[157]	26	60.27	35.54	2141.99	✓	×	-51.85	-6.87	-44.64	-48.44
[77]	18	27.06	17.45	472.19	✓	Considerably	-66.66	-58.19	-72.82	-88.63
Proposed SEUR-Latch	20	4.65	18.97	88.21	✓	Considerably	-62.96	-92.81	-70.45	-97.87
[58]	36	34.63	43.48	1505.71	✓	✓	-33.33	-46.49	-32.27	-63.76
[68]	63	6.54	57.31	374.8	✓	✓	+16.66	-89.89	-10.73	-90.97
Proposed DNUR-Latch	38	6.06	46.55	282.09	✓	✓	-29.63	-90.63	-27.49	-93.21

6.4.2 Performance Analysis of SEUR and DNUR NV-FF Designs

Thus far, the proposed radiation-hardened SEU and DNU resilient NV-FFs use the NV unit as a shadow latch [159] that is only active during the back-up and wake-up operations. This provides non-volatility to the FF designs without incurring energy overheads during the normal operation. Here, the performance of NV-FF designs that utilize the developed SEU-Resilient (SEUR) and DNU-Resilient (DNUR) latching circuits as the master latch and the developed NV-latch as slave latch as active components, i.e., incorporating for per cycle data backup, is also examined. Table 6.3 lists the structure of the NV-FF circuits investigated herein, as well as their vulnerable

nodes and sensitive node pairs. The SFC metric defined in Section 6.2 is used to evaluate the radiation-tolerance of the proposed NV-FFs, as listed in the last column of Table 6.3. The results indicate that the proposed SEU-resilient NV-FF circuits (SEUR-ShL and SEUR-Non-ShL) cannot tolerate $\sim 10\%$ of the DNUs, due to the existence of the unprotected sensitive node pairs, listed in the eighth column of the Table 6.3. Meanwhile, the DNU-resilient NV-FFs (DNUR-ShL and DNUR-Non-ShL) can tolerate 100% of single and double node upsets. However, this improvement in radiation tolerance is achieved at the cost of higher energy-consumption and area overhead as expressed in the Table 6.4. Moreover, it can be observed in Table 6.3 that the utilization of NV-latching circuitry as slave latch, i.e., active component for per cycle backup, or else the shadow latch does not have a significant impact on the SFC value.

Table 6.3: Radiation-tolerance analysis of the proposed NV-FF Designs. {ML=Master-Latch, SL=Slave-Latch, ShL=Shadow-Latch}

Design	NV-FF Structure			# Nodes	# Sensitive Nodes (n)	# Sensitive Node Pairs	# Unprotected Node Pairs	SFC (%) SEU/DNU
	ML	SL	ShL					
SEUR-ShL	SEUR	SEUR	NV	29	12: (n1-n8)(s0-s3)	66	6: (n1-n3, n4)(n5-n7, n8)	100/90.9
SEUR-Non-ShL	NV	SEUR	-	17	8: (n1-n4)(s0-s3)	28	3: (n1-n3, n4)	100/89.3
DNUR-ShL	DNUR	DNUR	NV	49	22: (n1-n18)(s0-s3)	231	0	100/100
DNUR-Non-ShL	NV	DNUR	-	26	13: (n1-n9)(s0-s3)	78	0	100/100

Table 6.4 lists the delay and power consumption of the NV-FFs investigated herein in different modes of operation. During the normal operation, i.e., latch/transfer mode, the delay and power consumption of the NV-FFs which use a shadow latch in their structure (SEUR-ShL and DNUR-ShL) designs are comparable to that of the conventional CMOS-based flip-flop, since the logic components utilized during normal operation are comparable to a CMOS flip-flop. In particular, the SEU-Resilient Shadow-Latch (SEUR-ShL) design uses one extra TG, and three additional TGs are utilized in the DNU-Resilient Shadow-Latch (DNUR-ShL) NV-FF circuit to isolate the CMOS-based datapath from that of NV logic circuits. During the back-up operation the logic state in the CMOS-based latch should be stored to the MTJ devices in NV shadow-latch circuit. Switching the state of the MTJs requires $\sim$ 98 fJ energy consumption and incurs $\sim$ 1.8 ns delay. During the

wake-up operation, the state of the MTJ is sensed by the developed DNU-resilient sense amplifier and its logic value is restored to the CMOS-based latch. Based on the results provided in Table 6.4, the energy consumption for the wake-up operation is approximately equal to 0.048 fJ and 0.074 fJ for the SEUR-ShL and DNUR-ShL NV-FF circuits, respectively. It is worth noting that, the back-up and wake-up operations in designs with Shadow-Latch (ShL) only occur when the FF is being power-gated.

Table 6.4: Performance comparison of the proposed NV-FFs.

Circuit Design		NV-FF SEUR-ShL	NV-FF SEUR-Non-ShL	NV-FF DNUR-ShL	NV-FF DNUR-Non-ShL
Device Count	MOS	63	35	105	53
	MTJ	2	2	2	2
Normal Mode	Power (μW)	24.56	135.52	52.42	162.31
	Delay (ns)	0.064	2.74	0.067	2.75
	Energy (fJ)	1.57	371.32	3.51	446.35
Back-up Mode	Power (μW)	54.77	NA	54.51	NA
	Delay (ns)	1.79	NA	1.8	NA
	Energy (fJ)	98.03	NA	98.12	NA
Wake-up Mode	Power (μW)	1.1	NA	1.58	NA
	Delay (ns)	0.044	NA	0.047	NA
	Energy (fJ)	0.048	NA	0.074	NA

In contrast, SEUR-Non-ShL and DNUR-Non-ShL designs do not have a shadow latch in their structure and the developed DNU-resilient NV-latching circuit is used as a master latch. Thus, the data should be written in and read out from the MTJ device in each cycle, which will significantly increase the power consumption and delay of the NV-FF's normal operation. The results listed in Table 6.4 show more than two orders of magnitude reduction in the energy consumption per cycle of the designs that use NV-element within a shadow latch instead of master latch, i.e. NV-FF SEUR-ShL and DNUR-ShL designs. This indicates that these designs incur roughly delay of 1.8 ns only before entering the standby mode. During the normal operation/mode, they incur negligible delay (64 and 67 ps, respectively). On the other hand, SEUR-Non-ShL and DNUR-Non-ShL designs backup the data bit in every clock cycle, and thus, they incur significant delay

(2.7 ns), which represents output delay of normal operations. Hence, although using a shadow latch incurs area overhead and adds back-up and wake-up operations to the NV-FF, it provides significant energy consumption advantages that makes the proposed SEU and DNU resilient NV-FFs suitable candidates to be used within energy-harvesting-powered systems, as well as intermittent and normally-off computing architectures.

6.5 Summary

In this chapter, energy-efficient and radiation-tolerant NV-FF circuits are developed using SHE-MTJ devices. First, a NV-latching circuit is designed that can tolerate multiple upset nodes using feedback transistors to discharge the radiation-induced charges injected to its sensitive nodes. Then, SEU and DNU resilient CMOS-based latching circuits are developed, which utilize interceptive feedback loops with the clock-gating Muller C-element at output stage to mask soft errors induced by the particles striking the sensitive nodes of the circuit. The developed radiation-hardened CMOS-based latching circuits realize more than 90% PDP improvement compared to conventional soft error resilient latching circuits.

Next, the designed latching circuits were leveraged to develop SEU and DNU resilient NV-FF circuits. In particular, the NV latching circuit is leveraged as a shadow latch to backup the data for the standby operation, while the CMOS-based latches are used as master and slave latches. The combination of the proposed DNU-resilient NV shadow latch and SEU-resilient CMOS-based master and slave latches achieved a radiation-hardened NV-FF, which could tolerate 100% of SEUs and more than 90% of DNUs. To realize 100% DNU-resiliency, the developed DNU-resilient latch circuit is leveraged as slave and master latches along with the DNU-tolerant NV shadow latch. Achievement of 100% DNU-tolerance is realized at the cost of higher area overhead and energy consumption compared to the proposed SEU-resilient NV-FF. In summary, the results obtained

quantify the clear trade-off between radiation immunity, area-overhead, and energy consumption. Consequently, the NV-FF designs developed herein can ensure correct operation in harsh environments in terms of radiation, while also supporting low-energy intermittent computing for IoT and space-based applications.

CHAPTER 7: CONCLUSION AND FUTURE WORK

This final chapter presents a summary of achievements of this dissertation and integration into state-of-the-art techniques that explore novel and bold ideas. In addition, the drawback(s) and inherent limitations of proposed techniques are discussed. Likewise, recommendations to improve circuit performance in terms of energy-efficiency and resilience of fault masking coverage are elaborated here. In the following sections, the technical summary of the proposed SE-mitigation approaches is presented and the possible future work is also discussed.

7.1 Technical Summary

The recent trends for the aggressive scaling of nanoscale CMOS device technologies have posed new challenges which significantly contribute to impact a system's performance and reliability. First, there exists an increased impact of Process Variation (PV) on device performance, while the second reliability challenge is an increased susceptibility of nanoscale device technology to soft errors. In particular, radiation-induced soft errors have become a key issue as they cause malfunctions in large-scale circuits and systems even for terrestrial applications. Similarly, the continuous shrinking of supply voltage reduces the required energy to induce transient and upset glitches, i.e., Single-Event Transient (SET) and Single-Event Upset (SEU), at the susceptible nodes of a logic circuit. In addition, operation at higher clock rate has also increased the probability of SETs to be captured by a latch/flip-flop. These reliability impacting factors have become major concern of recent device technology. Investigating these challenges underlying the design and evaluation of robust and energy-efficient modern computing systems for deeply-scaled technologies is the prime trends, prospects, and contributions of this dissertation. More importantly, how to improve robustness of circuits and systems towards soft errors while maintaining constricted energy bud-

get, how to reduce the static power dissipation, and how to tolerate the effects of process Variation (PV) are the major questions and aims that drive the research progress. Thus, addressing these challenges efficiently via careful designs, i.e., cost-effective, robust, and energy-efficient, were the major contributions of this dissertation. Since technology size scaling of CMOS devices has induced emerging issues that impact reliability of nanoscale logic circuits, it has been essential to address these reliability challenges at the design phase to improve the performance and resilience for the widespread use of VLSI systems towards these challenges.

In the first portion of this dissertation, a hybrid technique named Temporal Self-Voting Logic (TSVL), capturing benefits of both spatial and temporal redundancy approaches, has been introduced as an efficient approach that reduces the performance penalties of the spatial redundancy, i.e., conventional TMR, and improves the performance of temporal redundancy. The proposed technique is utilized to mask soft errors effects in logic paths, occurring due to a transient or an upset error. The developed TSVL latch employs a hybrid of spatial and temporal redundancy to reduce area overhead and power consumption of protection while ensuring the data integrity and improving systems' reliability. It utilizes a single datapath for combinational logic and duplicates the latches/flip-flops instead of duplicated/triplicated logic paths in the prior work. However, even though the design of TSVL technique reduces the area and power of the protection logic, it degrades the speed performance and increases complexity of the design.

On the other hand, near-threshold computing has been proposed as an effective strategy to reduce leakage power of deeply-scaled CMOS circuits. However, near-threshold strategies exacerbate the impact of delay variations on device performance and increase the susceptibility to soft errors due to narrow voltage margins, which makes it difficult to ensure correct performance of large-scale VLSI designs. Therefore, investigating the effects of PV on the SE-mitigation techniques at NTV region is the second portion of this dissertation. In particular, we provide detailed tradeoffs between energy consumption and resilience of fault masking coverage for redundancy-based SE-mitigation

approaches by introducing a new metric referred to as Fault Coverage Energy Ratio (FCER). Additionally, an important insight of this work is to identify redundancy-based hardening techniques that can deliver increased benefits for the leveraged tradeoffs within iso-energy constraints. Thus, the primary insight from this work is identification of redundancy-based hardening techniques that can deliver increased benefits in terms of the FCER for the leveraged tradeoffs within iso-energy constraints at NTV. In this study, we established the relationship between reliability-aware and energy-aware operation to improve reliability of low power computing systems in NTV region via the developed FCER metric, by slightly sacrificing fault resilience, i.e., tolerates few unmasked soft errors, while meeting the constrained budget of energy consumption. For instance, image processing applications are a promising field to employ our developed fault-tolerant approaches and FCER metric to immune mission-critical systems towards transient faults while delivering energy-efficiency and preserving high soft fault masking with an acceptable performance degradation. Finally, we demonstrated how the developed FCER metric provides a concise and effective optimization parameter for design of fault-tolerant circuits and systems, and its applicability to guide circuit synthesis to meet performance and robustness goals.

Other methods for increasing logic path reliability and realizing novel computing paradigms include asynchronous logic approaches [160–162], including spintronic-based [163] and QCA methods [164]. Thus, natural language systems constitute a growing area of future applications that can benefit from emerging computing approaches [165–167]. Recently, emerging spin-based devices are introduced as potential alternative and an intriguing candidate to alleviate leakage currents and scalability challenges of CMOS technology. Spintronic devices can provide the valuable attributes of non-volatility, near-zero leakage power, vertical device integration, and radiation hardness as a promising complement to CMOS technology. However, the immunity of read/write circuits towards radiation-induced soft errors needs to be adequately addressed. Therefore, in the last portion on this dissertation, radiation-immune hybrid CMOS/spintronic flip-flops/registers are designed

and evaluated for non-volatile applications. First, a spintronic NVFF circuit was developed using STT as a switching mechanism to write the data bit. Even though the proposed latching circuity achieved superior soft error resilience, i.e., significantly reduces the potential Single-Event Double Node Upset (SEDNU) rate, it still incurs high energy consumption due to the high current amplitude required for the switching. Thus, to reduce the protection overhead of extra logic and lower the switching energy, an energy-efficient SEU and DNU resilient latching circuits are introduced to be employed within logic datapaths to ensure data integrity. Here, instead of using the STT as switching mechanism, the SHE is employed to deliver less writing energy consumption while providing separate paths for read and write operations [11, 151]. Furthermore, the developed NVFF circuits are designed to work in four modes, including the following: Latch, Backup, Standby, and Wakeup mode. During normal operation (latching mode), the proposed SE-tolerant latching circuits will perform the same as CMOS-based FFs, while during the backup mode the data bit in the slave latch is stored in the NV Shadow-Latch (ShL). In the standby mode, the logic state is maintained in the NV elements, i.e., MTJs, and the circuit is power-gated with near-zero leakage power. Finally, during the wakeup mode, the stored logic state is sensed and recovered for normal latch operation. In this way, the design will incur comparable power consumption and delay compared to CMOS-based FFS, while alleviating the leakage power to be near-zero during the standby mode. Thus, the developed NVFFs can potentially assist in making the emerging spin-based latching circuits more feasible and resilient and also continuing the scalability process of CMOS VLSI chips, while delivering high-performance, energy-efficiency, and reliability for contemporary logic and memory architectures.

7.2 Future Work

The possible future work that can be investigated based on the work presented in this dissertation are highlighted below:

- One possible trends and agenda to extend this scholarly research is to deeply and seamlessly investigate the effect of soft-error in modern processor cores for nanoscale device technology. Evaluating resiliency of soft fault masking coverage for a single core processor towards Single-Event induced Multiple Upsets (SEMU) simultaneously can assist in designing reliable and efficient large-scale VLSI chips and also minimizing area overhead and power consumption of protection logic.

- Developing a deep learning classification method to accurately and efficiently estimate the vulnerable nodes, gates, and/or circuits towards transient faults of large-scale digital systems. This implies that the developed algorithm searches to identify the most vulnerable nodes in the design to accurately determine where to optimally allocate the spatial/temporal redundancy to improve the overall robustness of the targeted system. Identifying the most vulnerable nodes in a circuit, i.e., modular criticality, in a VLSI digital circuit turns to significantly improve robustness of large-scale CMOS designs, and thereby efficiently guiding the designer to realize minimized/limited performance penalties of area and energy overheads.

- Employing the Magnetic Electronic devices and materials instead of the heavy metal to reduce the write energy of NV latching circuits.

As Moore's Law continuing the scaling of CMOS device, additional challenges and reliability concerns will continue to arise regarding the unpredictable behavior of emerging large-scale circuits and systems, which makes it difficult to predict the performance of a VLSI design. In addition,

higher hardware complexity may experience new design challenges, and thus, reliability of VLSI systems will need further investigations. These will be a major concern of upcoming challenges to be identified, modeled, and addressed efficiently. Therefore, to deal with these issues, it will be essential to continue studying these research topics via thorough investigations to guide future research of emerging computing architectures at the ExaByte and ZettaByte scales.

APPENDIX: PERMISSION TO REUSE PUBLISHED MATERIAL

Title: Soft Error Effect Tolerant Temporal Self-Voting Checkers: Energy vs. Resilience Tradeoffs

Conference Proceedings: 2016 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)

Author: Faris S. Alghareb

Publisher: IEEE

Date: July 2016

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK CLOSE WINDOW

Title: Energy and Delay Tradeoffs of Soft-Error Masking for 16-nm FinFET Logic Paths: Survey and Impact of Process Variation in the Near-Threshold Region

Author: Faris S. Alghareb

Publication: Circuits and Systems Part II: Express Briefs, IEEE Transactions on

Publisher: IEEE

Date: June 2017

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

Title: Designing and Evaluating Redundancy-Based Soft-Error Masking on a Continuum of Energy versus Robustness

Author: Faris S. Alghareb

Publication: IEEE Transactions on Sustainable Computing

Publisher: IEEE

Date: 1 July-Sept. 2018

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

LIST OF REFERENCES

[1] S. Borkar, "Designing reliable systems from unreliable components: the challenges of transistor variability and degradation," *Micro, IEEE*, vol. 25, no. 6, pp. 10–16, Nov 2005.

[2] D. Mavis and P. Eaton, "Soft error rate mitigation techniques for modern microcircuits," in *Reliability Physics Symposium Proceedings, 2002. 40th Annual*, 2002, pp. 216–225.

[3] O. Amusan, L. Massengill, B. Bhuva, S. DasGupta, A. Witulski, and J. Ahlbin, "Design techniques to reduce SET pulse widths in deep-submicron combinational logic," *Nuclear Science, IEEE Transactions on*, vol. 54, no. 6, pp. 2060–2064, Dec 2007.

[4] J. Teifel, "Self-voting Dual-Modular-Redundancy circuits for single-event-transient mitigation," *Nuclear Science, IEEE Transactions on*, vol. 55, no. 6, pp. 3435–3439, Dec. 2008.

[5] D. Gomez Toro, M. Arzel, F. Seguin, and M. Jezequel, "Soft error detection and correction technique for radiation hardening based on C-element and BICS," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 61, no. 12, pp. 952–956, Dec. 2014.

[6] P. Reviriego, C. J. Bleakley, and J. A. Maestro, "Structural DMR: A technique for implementation of soft-error-tolerant FIR filters," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 58, no. 8, pp. 512–516, Aug. 2011.

[7] P. Hazucha, T. Karnik, J. Maiz, S. Walstra, B. Bloechel, J. Tschanz, G. Dermer, S. Hareland, P. Armstrong, and S. Borkar, "Neutron soft error rate measurements in a 90-nm CMOS process and scaling trends in SRAM from 0.25-/spl mu/m to 90-nm generation," in *Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International*, Dec 2003, pp. 21.5.1–21.5.4.

[8] H. Liu, M. Cotter, S. Datta, and V. Narayanan, "Soft-error performance evaluation on emerging low power devices," *Device and Materials Reliability, IEEE Transactions on*, vol. 14, no. 2, pp. 732–741, June 2014.

[9] M. Gadlage, J. Ahlbin, B. Narasimham, B. Bhuva, L. Massengill, R. Reed, R. Schrimpf, and G. Vizkelethy, "Scaling trends in SET pulse widths in sub-100 nm bulk CMOS processes," *Nuclear Science, IEEE Transactions on*, vol. 57, no. 6, pp. 3336–3341, Dec 2010.

[10] A. Dixit and A. Wood, "The impact of new technology on soft error rates," in *IEEE International Reliability Physics Symposium (IRPS)*, April 2011, pp. 5B.4.1–5B.4.7.

[11] S. Manipatruni, D. E. Nikonov, and I. A. Young, "Energy-delay performance of giant spin hall effect switching for dense magnetic memory," *Applied Physics Express*, vol. 7, no. 10, p. 103001, 2014.

[12] W. Kang, Z. Wang, Y. Zhang, J.-O. Klein, W. Lv, and W. Zhao, "Spintronic logic design methodology based on spin hall effect–driven magnetic tunnel junctions," *Journal of Physics D: Applied Physics*, vol. 49, no. 6, p. 065008, 2016.

[13] Y. Lakys, W. S. Zhao, J. O. Klein, and C. Chappert, "Hardening techniques for MRAM-based nonvolatile latches and logic," *IEEE Transactions on Nuclear Science*, vol. 59, no. 4, pp. 1136–1141, Aug 2012.

[14] Y. Lin, M. Zwolinski, and B. Halak, "A low-cost, radiation-hardened method for pipeline protection in microprocessors," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

[15] S. Mitra, N. Seifert, M. Zhang, Q. Shi, and K. S. Kim, "Robust system design with built-in soft-error resilience," *Computer*, vol. 38, no. 2, pp. 43–52, Feb 2005.

[16] D. Tang, C. He, Y. Li, H. Zang, C. Xiong, and J. Zhang, "Soft error reliability in advanced CMOS technologies-trends and challenges," *Science China Technological Sciences*, vol. 57, no. 9, pp. 1846–1857, Sep. 2014.

[17] N. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective*, 4th ed. USA: Addison-Wesley Publishing Company, 2010.

[18] R. Koga, S. H. Penzin, K. B. Crawford, and W. R. Crain, "Single event functional interrupt (SEFI) sensitivity in microcircuits," in *Radiation and Its Effects on Components and Systems, 1997. RADECS 97. Fourth European Conference on*, Sep 1997, pp. 311–318.

[19] I. Gomes, M. Martins, F. Lima Kastensmidt, A. Reis, R. Ribas, and S. Novales, "Methodology for achieving best trade-off of area and fault masking coverage in ATMR," in *Test Workshop - LATW, 2014 15th Latin American*, March 2014, pp. 1–6.

[20] M. Nicolaidis, "Design for soft error mitigation," *Device and Materials Reliability, IEEE Transactions on*, vol. 5, no. 3, pp. 405–418, Sept 2005.

[21] E. Cheng, S. Mirkhani, L. G. Szafaryn, C. Cher, H. Cho, K. Skadron, M. R. Stan, K. Lilja, J. A. Abraham, P. Bose, and S. Mitra, "Tolerating soft errors in processor cores using clear (cross-layer exploration for architecting resilience)," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1839–1852, Sept 2018.

[22] P. Dodd, M. Shaneyfelt, J. Schwank, and J. Felix, "Current and future challenges in radiation effects on CMOS electronics," *Nuclear Science, IEEE Transactions on*, vol. 57, no. 4, pp. 1747–1763, Aug 2010.

[23] A. Guar and H. Mahmoodi, "Impact of technology scaling on performance of domino logic in nano-scale CMOS," in *VLSI and System-on-Chip (VLSI-SoC), 2012 IEEE/IFIP 20th International Conference on*, Oct 2012, pp. 295–298.

[24] R. Sharpeshkar, *Ultra low power bioelectronics: Fundamentals, biomedical applications, and bio-inspired system*. Cambridge: Cambridge University Press, 2010.

[25] M. Omana, D. Rossi, and C. Metra, "High-performance robust latches," *IEEE Transactions on Computers*, vol. 59, no. 11, pp. 1455–1465, Nov 2010.

[26] V. Ferlet-Cavrois, L. Massengill, and P. Gouker, "Single event transients in digital CMOS–A review," *Nuclear Science, IEEE Transactions on*, vol. 60, no. 3, pp. 1767–1790, June 2013.

[27] N. Mahatme, N. Gaspard, T. Assis, S. Jagannathan, I. Chatterjee, T. Loveless, B. Bhuva, L. Massengill, S. Wen, and R. Wong, "Impact of technology scaling on the combinational logic soft error rate," in *Reliability Physics Symposium, 2014 IEEE International*, June 2014.

[28] T. Karnik and P. Hazucha, "Characterization of soft errors caused by single event upsets in CMOS processes," *Dependable and Secure Computing, IEEE Transactions on*, vol. 1, no. 2, pp. 128–143, April 2004.

[29] F. S. Alghareb, R. A. Ashraf, A. Alzahrani, and R. F. DeMara, "Energy and delay tradeoffs of soft-error masking for 16-nm FinFET logic paths: Survey and impact of process variation in the near-threshold region," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 64, no. 6, pp. 695–699, June 2017.

[30] B. Gaynor and S. Hassoun, "Fin shape impact on finfet leakage with application to multi-threshold and ultralow-leakage FinFET design," *Electron Devices, IEEE Transactions on*, vol. 61, no. 8, pp. 2738–2744, Aug. 2014.

[31] K. Kuhn, M. Giles, D. Becher, P. Kolar, A. Kornfeld, R. Kotlyar, S. Ma, A. Maheshwari, and S. Mudanai, "Process technology variation," *Electron Devices, IEEE Transactions on*, vol. 58, no. 8, pp. 2197–2208, Aug. 2011.

[32] H.-W. Su, Y. Li, Y.-Y. Chen, C.-Y. Chen, and H.-T. Chang, "Drain-induced-barrier lowering and subthreshold swing fluctuations in 16-nm-gate bulk FinFET devices induced by random discrete dopants," in *Device Research Conference (DRC), 2012 70th Annual*, June 2012.

[33] A. Brown, N. Daval, K. Bourdelle, B.-Y. Nguyen, and A. Asenov, "Comparative simulation analysis of process-induced variability in nanoscale soi and bulk trigate FinFETs," *Electron Devices, IEEE Transactions on*, vol. 60, no. 11, pp. 3611–3617, Nov 2013.

[34] N. Seifert, B. Gill, S. Jahinuzzaman, J. Basile, V. Ambrose, Q. Shi, R. Allmon, and A. Bramnik, "Soft error susceptibilities of 22 nm tri-gate devices," *Nuclear Science, IEEE Transactions on*, vol. 59, no. 6, pp. 2666–2673, Dec. 2012.

[35] R. Al-Haddad, R. S. Oreifej, R. Zand, A. Ejnioui, and R. F. DeMara, "Adaptive mitigation of radiation-induced errors and TDDB in reconfigurable logic fabrics," in *2015 IEEE 24th North Atlantic Test Workshop*, May 2015, pp. 23–32.

[36] K. Zhang, G. Bedette, and R. F. DeMara, "Triple modular redundancy with standby (TMRSB) supporting dynamic resource reconfiguration," in *2006 IEEE Autotestcon*, Sep. 2006, pp. 690–696.

[37] N. Imran and R. F. DeMara, "A self-configuring TMR scheme utilizing discrepancy resolution," in *2011 International Conference on Reconfigurable Computing and FPGAs*, Nov 2011, pp. 398–403.

[38] K. Zhang, R. F. DeMara, and C. A. Sharma, "Consensus-based evaluation for fault isolation and on-line evolutionary regeneration," in *Evolvable Systems: From Biology to Hardware*. Springer, 2005, pp. 12–24.

[39] Y. Tosaka, S. Satoh, T. Itakura, H. Ehara, T. Ueda, G. Woffinden, and S. Wender, "Measurement and analysis of neutron-induced soft errors in sub-half-micron CMOS circuits," *Electron Devices, IEEE Transactions on*, vol. 45, no. 7, pp. 1453–1458, Jul 1998.

[40] N. Cohen, T. S. Sriram, N. Leland, D. Moyer, S. Butler, and R. Flatley, "Soft error considerations for deep-submicron CMOS circuit applications," in *International Electron Devices Meeting 1999. Technical Digest (Cat. No.99CH36318)*, Dec 1999, pp. 315–318.

[41] Y. Tosaka, S. Satoh, and T. Itakura, "Neutron-induced soft error simulator and its accurate predictions," in *Simulation of Semiconductor Processes and Devices, 1997. SISPAD '97., 1997 International Conference on*, Sept 1997, pp. 253–256.

[42] S. Salehi and R. F. DeMara, "Process variation immune and energy aware sense amplifiers for resistive non-volatile memories," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2017, pp. 1–4.

[43] K.-M. Liu and C.-K. Lee, "Investigation of the random dopant fluctuations in 20-nm bulk MOSFETs and silicon-on-insulator FinFETs by ion implantation Monte Carlo simulation," in *Nanoelectronics Conference (INEC), 2013 IEEE 5th International*, Jan. 2013, pp. 263–266.

[44] Y. Li, H.-W. Su, Y.-Y. Chen, S.-C. Hsu, and W.-T. Huang, "The intrinsic parameter fluctuation on high-k/metal gate bulk FinFET devices," *Microelectronic Engineering*, vol. 109, pp. 302–305, 2013, insulating Films on Semiconductors.

[45] E. Amat, C. Almudever, N. Aymerich, R. Canal, and A. Rubio, "Impact of bulk/SOI 10nm FinFETs on 3T1D-DRAM cell performance," in *Solid-State and Integrated Circuit Technology (ICSICT), 2012 IEEE 11th International Conference on*, Oct 2012, pp. 1–3.

[46] Y. Ye, T. Liu, M. Chen, S. Nassif, and Y. Cao, "Statistical modeling and simulation of threshold variation under random dopant fluctuations and line-edge roughness," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, no. 6, pp. 987–996, June 2011.

[47] Y. Li, C.-H. Hwang, T.-Y. Li, and M.-H. Han, "Process-variation effect, metal-gate work-function fluctuation, and random-dopant fluctuation in emerging CMOS technologies," *Electron Devices, IEEE Transactions on*, vol. 57, no. 2, pp. 437–447, Feb 2010.

[48] P. Shivakumar, M. Kistler, S. Keckler, D. Burger, and L. Alvisi, "Modeling the effect of technology trends on the soft error rate of combinational logic," in *Dependable Systems and Networks, 2002. DSN 2002. Proceedings. International Conference on*, 2002, pp. 389–398.

[49] Q. Zhou and K. Mohanram, "Transistor sizing for radiation hardening," in *Reliability Physics Symposium Proceedings, 2004. 42nd Annual. 2004 IEEE International*, April 2004, pp. 310–315.

[50] B. Narasimham, B. L. Bhuva, R. D. Schrimpf, L. W. Massengill, M. J. Gadlage, O. A. Amusan, W. T. Holman, A. F. Witulski, W. H. Robinson, J. D. Black, J. M. Benedetto, and P. H. Eaton, "Characterization of digital single event transient pulse-widths in 130-nm and 90-nm CMOS technologies," *IEEE Transactions on Nuclear Science*, vol. 54, no. 6, pp. 2506–2511, Dec 2007.

[51] G. Bany Hamad, S. Hasan, O. Ait Mohamed, and Y. Savaria, "Abstracting single event transient characteristics variations due to input patterns and fan-out," in *Circuits and Systems (ISCAS), 2014 IEEE International Symposium on*, June 2014, pp. 1468–1471.

[52] G. Wirth, F. L. Kastensmidt, and I. Ribeiro, "Single event transients in logic circuits-load and propagation induced pulse broadening," *Nuclear Science, IEEE Transactions on*, vol. 55, no. 6, pp. 2928–2935, Dec 2008.

[53] Y. Li, M. Li, J. Schrimpf, and D. Fleetwood, "Characterizing, modeling, and simulating soft error susceptibility in cell-based designs in highly scaled technologies," in *Solid-State and Integrated Circuit Technology (ICSICT), 2012 IEEE 11th International Conference on*, Oct 2012, pp. 1–6.

[54] V. Ferlet-Cavrois, P. Paillet, M. Gaillardin, D. Lambert, J. Baggio, J. Schwank, G. Vizkelethy, M. Shaneyfelt, K. Hirose, E. Blackmore, O. Faynot, C. Jahan, and L. Tosti, "Statistical analysis of the charge collected in SOI and bulk devices under heavy lon and proton irradiation mdash-implications for digital SETs," *Nuclear Science, IEEE Transactions on*, vol. 53, no. 6, pp. 3242–3252, Dec 2006.

[55] H. Nakamura, T. Uemura, K. Takeuchi, T. Fukuda, S. Kumashiro, and T. Mogami, "Scaling effect and circuit type dependence of neutron induced single event transient," in *Reliability Physics Symposium (IRPS), 2012 IEEE International*, April 2012, pp. 3C.3.1–3C.3.7.

[56] R. A. Ashraf, R. Gioiosa, G. Kestor, R. F. DeMara, C. Y. Cher, and P. Bose, "Understanding the propagation of transient errors in HPC applications," in *SC15: International Conference for High Performance Computing, Networking, Storage and Analysis*, Nov 2015, pp. 1–12.

[57] J. Yao, Z. Ye, M. Li, Y. Li, R. Schrimpf, D. Fleetwood, and Y. Wang, "Statistical analysis of soft error rate in digital logic design including process variations," *Nuclear Science, IEEE Transactions on*, vol. 59, no. 6, pp. 2811–2817, Dec 2012.

[58] K. Katsarou and Y. Tsiatouhas, "Double node charge sharing SEU tolerant latch design," in *2014 IEEE 20th International On-Line Testing Symposium (IOLTS)*, July 2014, pp. 122–127.

[59] B. D. Sierawski, B. L. Bhuva, and L. W. Massengill, "Reducing soft error rate in logic circuits through approximate logic functions," *IEEE Transactions on Nuclear Science*, vol. 53, no. 6, pp. 3417–3421, Dec 2006.

[60] R. Rao, D. Blaauw, and D. Sylvester, "Soft error reduction in combinational logic using gate resizing and flipflop selection," in *Computer-Aided Design, 2006. ICCAD '06. IEEE/ACM International Conference on*, Nov 2006, pp. 502–509.

[61] J. Chen, S. Chen, Y. He, J. Qin, B. Liang, B. Liu, and P. Huang, "Novel layout technique for single-event transient mitigation using dummy transistor," *Device and Materials Reliability, IEEE Transactions on*, vol. 13, no. 1, pp. 177–184, March 2013.

[62] J. Ahlbin, M. Gadlage, D. Ball, A. Witulski, B. Bhuva, R. Reed, G. Vizkelethy, and L. Massengill, "The effect of layout topology on single-event transient pulse quenching in a 65 nm bulk CMOS process," *Nuclear Science, IEEE Transactions on*, vol. 57, no. 6, pp. 3380–3385, Dec 2010.

[63] L. H.-H. Kelin, L. Klas, B. Mounaim, R. Prasanthi, I. R. Linscott, U. S. Inan, and M. Subhasish, "LEAP: Layout design through error-aware transistor positioning for soft-error resilient sequential cell design," in *2010 IEEE International Reliability Physics Symposium*, May 2010, pp. 203–212.

[64] P. Roche, F. Jacquet, C. Caillat, and J.-P. Schoellkopf, "An alpha immune and ultra low neutron SER high density SRAM," in *Reliability Physics Symposium Proceedings, 2004. 42nd Annual. 2004 IEEE International*, April 2004, pp. 671–672.

[65] M. Casey, B. Bhuva, J. Black, and L. Massengill, "HBD using cascode-voltage switch logic gates for SET tolerant digital designs," *Nuclear Science, IEEE Transactions on*, vol. 52, no. 6, pp. 2510–2515, Dec 2005.

[66] D. Limbrick, D. Black, K. Dick, N. Atkinson, N. Gaspard, J. Black, W. Robinson, and A. Witulski, "Impact of logic synthesis on soft error vulnerability using a 90-nm bulk CMOS digital cell library," in *Southeastcon, 2011 Proceedings of IEEE*, March 2011, pp. 430–434.

[67] F. S. Alghareb, M. Lin, and R. F. DeMara, "Soft error effect tolerant temporal self-voting checkers: Energy vs. resilience tradeoffs," in *2016 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, July 2016, pp. 571–576.

[68] A. Yan, Z. Huang, M. Yi, X. Xu, Y. Ouyang, and H. Liang, "Double-node-upset-resilient latch design for nanoscale CMOS technology," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 6, pp. 1978–1982, June 2017.

[69] F. S. Alghareb, R. A. Ashraf, and R. F. DeMara, "Designing and evaluating redundancy-based soft-error masking on a continuum of energy versus robustness," *IEEE Transactions on Sustainable Computing*, vol. 3, no. 3, pp. 139–152, July 2018.

[70] D. Ernst, N. S. Kim, S. Das, S. Pant, R. Rao, T. Pham, C. Ziesler, D. Blaauw, T. Austin, K. Flautner, and T. Mudge, "Razor: a low-power pipeline based on circuit-level timing speculation," in *Microarchitecture, 2003. MICRO-36. Proceedings. 36th Annual IEEE/ACM International Symposium on*, Dec 2003, pp. 7–18.

[71] S. Jagannathan, T. D. Loveless, B. L. Bhuva, S. J. Wen, R. Wong, M. Sachdev, D. Rennie, and L. W. Massengill, "Single-event tolerant flip-flop design in 40-nm bulk cmos technology," *IEEE Transactions on Nuclear Science*, vol. 58, no. 6, pp. 3033–3037, Dec 2011.

[72] N. Seifert, V. Ambrose, B. Gill, Q. Shi, R. Allmon, C. Recchia, S. Mukherjee, N. Nassif, J. Krause, J. Pickholtz, and A. Balasubramanian, "On the radiation-induced soft error performance of hardened sequential elements in advanced bulk CMOS technologies," in *2010 IEEE International Reliability Physics Symposium*, May 2010, pp. 188–197.

[73] A. Watkins and S. Tragoudas, "Radiation hardened latch designs for double and triple node upsets," *IEEE Transactions on Emerging Topics in Computing*, pp. 1–1, 2018.

[74] T. Calin, M. Nicolaidis, and R. Velazco, "Upset hardened memory design for submicron CMOS technology," *Nuclear Science, IEEE Transactions on*, vol. 43, no. 6, pp. 2874–2878, Dec 1996.

[75] T. D. Loveless, S. Jagannathan, T. Reece, J. Chetia, B. L. Bhuva, M. W. McCurdy, L. W. Massengill, S. J. Wen, R. Wong, and D. Rennie, "Neutron- and proton-induced single event upsets for D- and DICE-flip/flop designs at a 40 nm technology node," *IEEE Transactions on Nuclear Science*, vol. 58, no. 3, pp. 1008–1014, June 2011.

[76] M. D'Alessio, M. Ottavi, and F. Lombardi, "Design of a nanometric CMOS memory cell for hardening to a single event with a multiple-node upset," *IEEE Transactions on Device and Materials Reliability*, vol. 14, no. 1, pp. 127–132, March 2014.

[77] H. B. Wang, Y. Q. Li, L. Chen, L. X. Li, R. Liu, S. Baeg, N. Mahatme, B. L. Bhuva, S. J. Wen, R. Wong, and R. Fung, "An SEU-tolerant DICE latch design with feedback transistors," *IEEE Transactions on Nuclear Science*, vol. 62, no. 2, pp. 548–554, April 2015.

[78] P. Reviriego, O. Ruano, M. Flanagan, S. Pontarelli, and J. Maestro, "An efficient technique to protect serial shift registers against soft errors," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 60, no. 8, pp. 512–516, Aug. 2013.

[79] R. Al-Haddad, R. Oreifej, R. Ashraf, and R. F. DeMara, "Sustainable modular adaptive redundancy technique emphasizing partial reconfiguration for reduced power consumption," *International Journal of Reconfigurable Computing*, vol. 2011, no. 430808, 2011.

[80] R. S. Oreifej, R. Al-Haddad, R. Zand, R. A. Ashraf, and R. F. DeMara, "Survivability modeling and resource planning for self-repairing reconfigurable device fabrics," *IEEE Transactions on Cybernetics*, vol. 48, no. 2, pp. 780–792, Feb 2018.

[81] R. S. Oreifej, R. Al-Haddad, R. A. Ashraf, and R. F. DeMara, "Sustainability assurance modeling for SRAM-based FPGA evolutionary self-repair," in *2014 IEEE International Conference on Evolvable Systems*, Dec 2014, pp. 17–22.

[82] T. R. Weatherford, "Radiation effects in high speed III-V integrated circuits," *International journal of high speed electronics and systems*, vol. 13, no. 01, pp. 277–292, 2003.

[83] J. Celis, S. De La Rosa Nieves, C. Fuentes, S. Gutierrez, and A. Saenz-Otero, "Methodology for designing highly reliable fault tolerance space systems based on COTS devices," in *Proceedings of the IEEE International Systems Conference (SysCon)*, 2013, pp. 591–594.

[84] C. Engelmann, H. Ong, and S. L. Scott, "The case for modular redundancy in large-scale high performance computing systems," in *Proceedings of the IASTED International Conference on Parallel and Distributed Computing and Networks (PDCN)*, vol. 641, Feb 2009, pp. 189–194.

[85] N. Avirneni and A. Somani, "Low overhead soft error mitigation techniques for high-performance and aggressive designs," *Computers, IEEE Transactions on*, vol. 61, no. 4, pp. 488–501, April 2012.

[86] N. Imran and R. DeMara, "Heterogeneous concurrent error detection (hCED) based on output anticipation," in *Reconfigurable Computing and FPGAs (ReConFig), 2011 International Conference on*, Nov. 2011, pp. 61–66.

[87] P. Reviriego, C. Bleakley, and J. Maestro, "Diverse double modular redundancy: A new direction for soft-error detection and correction," *Design Test, IEEE*, vol. 30, no. 2, pp. 87–95, April 2013.

[88] K. Salako and L. Strigini, “When does "diversity" in development reduce common failures? insights from probabilistic modeling,” *IEEE Transactions on Dependable and Secure Computing*, vol. 11, no. 2, pp. 193–206, March 2014.

[89] S. Mitra and E. McCluskey, “Combinational logic synthesis for diversity in duplex systems,” in *Test Conference, 2000. Proceedings. International*, 2000, pp. 179–188.

[90] R. A. Ashraf, O. Mouri, R. Jadaa, and R. F. DeMara, “Design-for-diversity for improved fault-tolerance of TMR systems on FPGAs,” in *2011 International Conference on Reconfigurable Computing and FPGAs*, Nov 2011, pp. 99–104.

[91] K. Wu and R. Karri, “Algorithm level recomputing using allocation diversity: a register transfer level approach to time redundancy-based concurrent error detection,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 21, no. 9, pp. 1077–1087, Sep. 2002.

[92] S. Mirkhani, S. Mitra, C. Y. Cher, and J. Abraham, “Efficient soft error vulnerability estimation of complex designs,” in *2015 Design, Automation Test in Europe Conference Exhibition (DATE)*, March 2015, pp. 103–108.

[93] N. Seifert, P. Slankard, M. Kirsch, B. Narasimham, V. Zia, C. Brookreson, A. Vo, S. Mitra, B. Gill, and J. Maiz, “Radiation-induced soft error rates of advanced CMOS bulk devices,” in *2006 IEEE International Reliability Physics Symposium Proceedings*, March 2006, pp. 217–225.

[94] D. Chabi, W. Zhao, E. Deng, Y. Zhang, N. B. Romdhane, J. O. Klein, and C. Chappert, “Ultra low power magnetic flip-flop based on checkpointing/power gating and self-enable mechanisms,” *IEEE Trans on Circuits and Systems I*, vol. 61, no. 6, pp. 1755–1765, June 2014.

[95] N. Eftaxiopoulos, N. Axelos, and K. Pekmestzi, "DONUT: A double node upset tolerant latch," in *VLSI (ISVLSI), 2015 IEEE Computer Society Annual Symposium on*, July 2015, pp. 509–514.

[96] S. Mukherjee, *Architecture design for soft errors*, 1st ed. Morgan Kaufmann Publishers, San Francisco, CA, 2008.

[97] J. L. Hennessy and D. A. Patterson, *Computer architecture: a quantitative approach*, 4th ed. Morgan Kaufmann Publishers, San Francisco, CA, 2006.

[98] D. Lipetz and E. Schwarz, "Self checking in current floating-point units," in *2011 IEEE 20th Sympo. on Computer Arith.*, July 2011, pp. 73–76.

[99] S. Raasch, A. Biswas, J. Stephan, P. Racunas, and J. Emer, "A fast and accurate analytical technique to compute the AVF of sequential bits in a processor," in *Proceedings of the 48th International Symposium on Microarchitecture*, ser. MICRO-48. New York, NY, USA: ACM, 2015, pp. 738–749. [Online]. Available: http://doi.acm.org/10.1145/2830772.2830829

[100] C. Bolchini, L. Pomante, F. Salice, and D. Sciuto, "The design of reliable devices for mission-critical applications," *IEEE Trans. on Instrumentation and Measurement*, vol. 52, no. 6, pp. 1703–1712, Dec 2003.

[101] M. Pourali, "Incorporating common cause failures in mission-critical facilities reliability analysis," *IEEE Transactions on Industry Applications*, vol. 50, no. 4, pp. 2883–2890, July 2014.

[102] F. S. Alghareb and R. F. DeMara, "Design and evaluation of DNU-tolerant registers for resilient architectural state storage," in *2019 the 28th ACM Great Lakes Symposium on VLSI (GLSVLSI)*, May 2019.

[103] B. Shim, S. R. Sridhara, and N. R. Shanbhag, "Reliable low-power digital signal processing via reduced precision redundancy," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 12, no. 5, pp. 497–510, May 2004.

[104] R. Travessini, P. R. C. Villa, F. L. Vargas, and E. A. Bezerra, "Processor core profiling for SEU effect analysis," in *2018 IEEE 19th Latin-American Test Symposium (LATS)*, March 2018, pp. 1–6.

[105] X. Iturbe, B. Venu, and E. Ozer, "Soft error vulnerability assessment of the real-time safety-related ARM Cortex-R5 CPU," in *2016 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT)*, Sept 2016, pp. 91–96.

[106] R. Ashraf, R. Gioiosa, and R. F. DeMara, "Exploring the effect of compiler optimizations on the reliability of HPC applications," in *IEEE Workshop on Dependable Parallel, Distributed and Network-Centric Systems (DPDNS)*, June 2017.

[107] K. Johansson, P. Dyreklev, B. Granbom, M. Calver, S. Fourtine, and O. Feuillatre, "In-flight and ground testing of single event upset sensitivity in static RAMs," *Nuclear Science, IEEE Transactions on*, vol. 45, no. 3, pp. 1628–1632, Jun 1998.

[108] N. Khoshavi, , , and R. F. DeMara, "Bit-upset vulnerability factor for eDRAM last level cache immunity analysis," in *2016 17th International Symposium on Quality Electronic Design (ISQED)*, March 2016, pp. 6–11.

[109] C. Slayman, "Soft error trends and mitigation techniques in memory devices," in *Reliability and Maintainability Symposium (RAMS), 2011 Proceedings - Annual*, Jan 2011, pp. 1–5.

[110] G. Schoof, M. Methfessel, and R. Kraemer, "Fault-tolerant ASIC design for high system dependability," in *Advanced Microsystems for Automotive Applications 2009*. Springer, 2009, pp. 369–382.

[111] H. Fuketa, R. Harada, M. Hashimoto, and T. Onoye, "Measurement and analysis of alpha-particle-induced soft errors and multiple-cell upsets in 10T subthreshold SRAM," *Device and Materials Reliability, IEEE Transactions on*, vol. 14, no. 1, pp. 463–470, March 2014.

[112] F. Ruckerbauer and G. Georgakos, "Soft error rates in 65nm SRAMs-analysis of new phenomena," in *On-Line Testing Symposium, 2007. IOLTS 07. 13th IEEE International*, July 2007, pp. 203–204.

[113] B. Liu, S. Chen, B. Liang, Z. Liu, and Z. Zhao, "Temperature dependency of charge sharing and MBU sensitivity in 130-nm CMOS technology," *Nuclear Science, IEEE Transactions on*, vol. 56, no. 4, pp. 2473–2479, Aug 2009.

[114] E. Ibe, S. Chung, S. Wen, H. Yamaguchi, Y. Yahagi, H. Kameyama, S. Yamamoto, and T. Akioka, "Spreading diversity in multi-cell neutron-induced upsets with device scaling," in *Custom Integrated Circuits Conference, 2006. CICC '06. IEEE*, Sept 2006, pp. 437–444.

[115] N. Mahatme, B. Bhuva, N. Gaspard, T. Assis, Y. Xu, P. Marcoux, M. Vilchis, B. Narasimham, A. Shih, S.-J. Wen, R. Wong, N. Tam, M. Shroff, S. Koyoma, and A. Oates, "Terrestrial SER characterization for nanoscale technologies: A comparative study," in *Reliability Physics Symposium (IRPS), 2015 IEEE International*, April 2015, pp. 4B.4.1–4B.4.7.

[116] O. Amusan, A. Steinberg, A. Witulski, B. Bhuva, J. Black, M. Baze, and L. Massengill, "Single event upsets in a 130 nm hardened latch design due to charge sharing," in *Reliability physics symposium, 2007. proceedings. 45th annual. ieee international*, April 2007, pp. 306–311.

[117] M. Nicolaidis, "Design techniques for soft-error mitigation," in *IC Design and Technology (ICICDT), 2010 IEEE International Conference on*, June 2010, pp. 208–214.

[118] V. Petrovic, G. Schoof, and Z. Stamenkovic, "Fault-tolerant TMR and DMR circuits with latchup protection switches," *Microelectronics Reliability*, vol. 54, no. 8, pp. 1613 – 1626, 2014.

[119] N. Rollins, M. Wirthlin, M. Caffrey, and P. Graham, "Evaluating TMR techniques in the presence of single event upsets," in *Proceedings fo the 6th Annual International Conference on Military and Aerospace Programmable Logic Devices (MAPLD). Washington, DC: NASA Office of Logic Design, AIAA*, 2003, p. P63.

[120] R.-M. Huang and C.-P. Wen, "Advanced soft-error-rate (SER) estimation with striking-time and multi-cycle effects," in *Design Automation Conference (DAC), 2014 51st ACM/EDAC/IEEE*, June 2014, pp. 1–6.

[121] N. Mahatme, S. Jagannathan, T. Loveless, L. Massengill, B. Bhuva, S.-J. Wen, and R. Wong, "Comparison of combinational and sequential error rates for a deep submicron process," *Nuclear Science, IEEE Transactions on*, vol. 58, no. 6, pp. 2719–2725, Dec 2011.

[122] N. Pinckney, D. Blaauw, and D. Sylvester, "Low-power near-threshold design: Techniques to improve energy efficiency," *Solid-State Circuits Magazine, IEEE*, vol. 7, no. 2, pp. 49–57, Spring 2015.

[123] R. Dreslinski, M. Wieckowski, D. Blaauw, D. Sylvester, and T. Mudge, "Near-threshold computing: Reclaiming Moore's law through energy efficient integrated circuits," *Proceedings of the IEEE*, vol. 98, no. 2, pp. 253–266, 2010.

[124] M. Snir, R. W. Wisniewski, J. A. Abraham, S. V. Adve, S. Bagchi, P. Balaji, J. Belak, P. Bose, F. Cappello, B. Carlson, A. A. Chien, P. Coteus, N. A. Debardeleben, P. Diniz, C. Engelmann, M. Erez, S. Fazzari, A. Geist, R. Gupta, F. Johnson, S. Krishnamoorthy, S. Leyffer, D. Liberty, S. Mitra, T. S. Munson, R. Schreiber, J. Stearley, and E. V. Hensbergen,

"Addressing failures in exascale computing," *International Journal of High Performance Computing*, vol. ANL/MCS-P5022-0913, Mar 2013.

[125] Y.-P. Fang and A. Oates, "Neutron-induced charge collection simulation of bulk FinFET SRAMs compared with conventional planar SRAMs," *Device and Materials Reliability, IEEE Transactions on*, vol. 11, no. 4, pp. 551–554, Dec. 2011.

[126] E. Amat, C. Almudever, N. Aymerich, R. Canal, and A. Rubio, "Suitability of the FinFET 3T1D cell beyond 10 nm," *Nanotechnology, IEEE Transactions on*, vol. 13, no. 5, pp. 926–932, Sept 2014.

[127] M. Shokrolah-Shirazi and S. Miremadi, "FPGA-based fault injection into synthesizable verilog HDL models," in *Secure System Integration and Reliability Improvement, 2008. SSIRI '08. Second International Conference on*, July 2008, pp. 143–149.

[128] E. Krimer, R. Pawlowski, M. Erez, and P. Chiang, "Synctium: a near-threshold stream processor for energy-constrained parallel applications," *IEEE Computer Architecture Letters*, vol. 9, no. 1, pp. 21–24, 2010.

[129] H. Kaul, M. Anders, S. Hsu, A. Agarwal, R. Krishnamurthy, and S. Borkar, "Near-threshold voltage (NTV) design: Opportunities and challenges," in *Proceedings of the 49th ACM/EDAC/IEEE Annual Design Automation Conference (DAC)*, June 2012, pp. 1153–1158.

[130] J. Torrellas, "Extreme-scale computer architecture: Energy efficiency from the ground up," in *Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014*, March 2014, pp. 1–5.

[131] W. Zhao and Y. Cao, "New generation of Predictive Technology Model for sub-45nm design exploration," in *Quality Electronic Design, 2006. ISQED '06. 7th International Symposium on*, March 2006.

[132] M. Gadlage, J. Ahlbin, B. Bhuva, N. Hooten, N. Dodds, R. Reed, L. Massengill, R. Schrimpf, and G. Vizkelethy, "Alpha-particle and focused-ion-beam-induced single-event transient measurements in a bulk 65-nm CMOS technology," *Nuclear Science, IEEE Transactions on*, vol. 58, no. 3, pp. 1093–1097, June 2011.

[133] F. S. Alghareb, R. Zand, and R. F. DeMara, "High-performance double node upset-tolerant non-volatile flip-flop design," in *2018 IEEE SoutheastCon (SECon-2018)*, April 2018, pp. 1–6.

[134] W. Zhao, E. Deng, J.-O. Klein, Y. Cheng, Y. Zhang, C. Chappert *et al.*, "A radiation hardened hybrid spintronic/cmos nonvolatile unit using magnetic tunnel junctions," *Journal of Physics D: Applied Physics*, vol. 47, no. 40, p. 405003, 2014.

[135] Y. Zhang, W. Zhao, Y. Lakys, J. O. Klein, J. V. Kim, D. Ravelosona, and C. Chappert, "Compact modeling of perpendicular-anisotropy CoFeB/MgO magnetic tunnel junctions," *IEEE Transactions on Electron Devices*, vol. 59, no. 3, pp. 819–826, March 2012.

[136] Y. Liu, Z. Li, H. Li, Y. Wang, X. Li, K. Ma, S. Li, M. F. Chang, S. John, Y. Xie, J. Shu, and H. Yang, "Ambient energy harvesting nonvolatile processors: From circuit to system," in *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, June 2015, pp. 1–6.

[137] S. D. Pyle, H. Li, and R. F. DeMara, "Compact low-power instant store and restore D flip-flop using a self-complementing spintronic device," *Electronics Letters*, vol. 52, no. 14, pp. 1238–1240, 2016.

[138] S. Gerardin and A. Paccagnella, "Present and future non-volatile memories for space," *IEEE Transactions on Nuclear Science*, vol. 57, no. 6, pp. 3016–3039, Dec 2010.

[139] S. Salehi, D. Fan, and R. F. DeMara, "Survey of STT-MRAM cell design strategies: Taxonomy and sense amplifier tradeoffs for resiliency," *J. Emerg. Technol. Comput. Syst.*, vol. 13, no. 3, pp. 48:1–48:16, Apr 2017.

[140] R. Rajaei, S. B. Mamaghani, and F. Eslaminasab, "Radiation hardening by design for non-volatile magnetic flip-flops," in *The 1st International Conference on New Research Achievements in Electrical and Computer Engineering, Tehran, Iran*, 2016.

[141] W. Zhao, E. Belhaire, C. Chappert, F. Jacquet, and P. Mazoyer, "New non-volatile logic based on spin-MTJ," *physica status solidi (a)*, vol. 205, no. 6, pp. 1373–1377, 2008.

[142] B. Behin-Aein, J.-P. Wang, and R. Wiesendanger, "Computing with spins and magnets," *MRS Bulletin*, vol. 39, no. 08, pp. 696–702, 2014.

[143] R. Zand, A. Roohi, D. Fan, and R. F. DeMara, "Energy-efficient nonvolatile reconfigurable logic using spin hall effect-based lookup tables," *IEEE Transactions on Nanotechnology*, vol. 16, no. 1, pp. 32–43, Jan 2017.

[144] J. C. Slonczewski, "Current-driven excitation of magnetic multilayers," *Journal of Magnetism and Magnetic Materials*, vol. 159, no. 1, pp. L1–L7, 1996.

[145] W. Kang, Y. Zhang, Z. Wang, J.-O. Klein, C. Chappert, D. Ravelosona, G. Wang, Y. Zhang, and W. Zhao, "Spintronics: Emerging ultra-low-power circuits and systems beyond MOS technology," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 12, no. 2, p. 16, 2015.

[146] Y. Suge, Z. Xiaolin, Z. Yuanfu, L. Lin, and W. Hanning, "Modeling and simulation of single-event effect in CMOS circuit," *Journal of Semiconductors*, vol. 36, no. 11, p. 111002, 2015.

[147] M. Omana, D.Rossi, and C. Metra, "Latch susceptibility to transient faults and new hardening approach," *IEEE Transactions on Computers*, vol. 56, no. 9, pp. 1255–1268, Sept 2007.

[148] D. Chabi, W. Zhao, J. O. Klein, and C. Chappert, "Design and analysis of radiation hardened sensing circuits for spin transfer torque magnetic memory and logic," *IEEE Transactions on Nuclear Science*, vol. 61, no. 6, pp. 3258–3264, Dec 2014.

[149] K. Katsarou and Y. Tsiatouhas, "Soft error interception latch: double node charge sharing SEU tolerant design," *Electronics Letters*, vol. 51, no. 4, pp. 330–332, 2015.

[150] F. S. Alghareb, R. Zand, and R. F. DeMara, "Non-volatile spintronic flip-flop design for energy-efficient SEU and DNU resilience," *IEEE Transactions on Magnetics*, vol. 55, no. 3, pp. 1–11, March 2019.

[151] L. Liu, T. Moriyama, D. C. Ralph, and R. A. Buhrman, "Spin-torque ferromagnetic resonance induced by the spin hall effect," *Phys. Rev. Lett.*, vol. 106, p. 036601, Jan 2011.

[152] A. Roohi, R. Zand, D. Fan, and R. F. DeMara, "Voltage-based concatenatable full adder using spin hall effect switching," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. PP, no. 99, pp. 1–1, 2017.

[153] J. Guo, L. Xiao, and Z. Mao, "Novel low-power and highly reliable radiation hardened memory cell for 65 nm CMOS technology," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 7, pp. 1994–2001, July 2014.

[154] R. Zand, A. Roohi, and R. F. DeMara, "Energy-efficient and process-variation-resilient write circuit schemes for spin hall effect MRAM device," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 9, pp. 2394–2401, Sept 2017.

[155] X. Fong, Y. Kim, K. Yogendra, D. Fan, A. Sengupta, A. Raghunathan, and K. Roy, "Spin-transfer torque devices for logic and memory: Prospects and perspectives," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 35, no. 1, pp. 1–22, Jan 2016.

[156] R. Zand and R. F. DeMara, "Radiation-hardened MRAM-based LUT for non-volatile FPGA soft error mitigation with multi-node upset tolerance," *Journal of Physics D: Applied Physics*, vol. 50, no. 50, p. 505002, 2017.

[157] A. Yan, H. Liang, Z. Huang, C. Jiang, Y. Ouyang, and X. Li, "An SEU resilient, SET filterable and cost effective latch in presence of PVT variations," *Microelectronics Reliability*, vol. 63, no. Supplement C, pp. 239 – 250, 2016.

[158] M. Fazeli, S. G. Miremadi, A. Ejlali, and A. Patooghy, "Low energy single event upset/single event transient-tolerant latch for deep submicron technologies," *IET Computers Digital Techniques*, vol. 3, no. 3, pp. 289–303, May 2009.

[159] R. Bishnoi, F. Oboril, and M. B. Tahoori, "Design of defect and fault-tolerant nonvolatile spintronic flip-flops," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 4, pp. 1421–1432, April 2017.

[160] W. Kuang, P. Zhao, J. S. Yuan, and R. F. DeMara, "Design of asynchronous circuits for high soft error tolerance in deep submicrometer CMOS circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 18, no. 3, pp. 410–422, March 2010.

[161] S. Smith, R. DeMara, J. Yuan, D. Ferguson, and D. Lamb, "Optimization of null convention self-timed circuits," *Integration*, vol. 37, no. 3, pp. 135 – 165, 2004.

[162] S. Smith, R. DeMara, J. Yuan, M. Hagedorn, and D. Ferguson, "Delay-insensitive gate-level pipelining," *Integration*, vol. 30, no. 2, pp. 103 – 131, 2001.

[163] Y. Bai, R. F. DeMara, J. Di, and M. Lin, "Clockless spintronic logic: A robust and ultra-low power computing paradigm," *IEEE Transactions on Computers*, vol. 67, no. 5, pp. 631–645, May 2018.

[164] A. M. Chabi, A. Roohi, H. Khademolhosseini, S. Sheikhfaal, S. Angizi, K. Navi, and R. F. DeMara, "Towards ultra-efficient QCA reversible circuits," *Microprocessors and Microsystems*, vol. 49, pp. 127 – 138, 2017.

[165] V. Hung, A. Gonzalez, and R. DeMara, "Towards a context-based dialog management layer for expert systems," in *2009 International Conference on Information, Process, and Knowledge Management*, Feb 2009, pp. 60–65.

[166] V. Hung, M. Elvir, A. Gonzalez, and R. DeMara, "Towards a method for evaluating naturalness in conversational dialog systems," in *2009 IEEE International Conference on Systems, Man and Cybernetics*, Oct 2009, pp. 1236–1241.

[167] R. F. DeMara and D. I. Moldovan, "Performance indices for parallel marker-propagation," in *1991 International Conference on Parallel Processing (ICPP)*, Aug 1991, pp. 658–659.